

The constant gardeners

While pottering away in a garden near you, botanists are playing an increasingly sophisticated role in studying plant diversity. They should continue to broaden their scientific reach.

Thank goodness for botanical gardens. Without them, who would compile the flora of Mesoamerica? Who would identify a rare member of the Melastomataceae family, or determine the most likely pollinator for a newly discovered orchid? Imagine the intellectual poverty of a world in which no plant researcher studied anything but the model organism *Arabidopsis thaliana*, an unprepossessing mustard cress.

Universities do still house some taxonomists and whole-organism biologists. But as they retire, they are ever more likely to be replaced with geneticists or a similarly fashionable brand of molecular biologist, and their herbaria shunted off to the nearest botanical garden. "The number of well-trained taxonomists is shrinking," warns one leading systemic biologist. "They are a dying breed."

Not that the research interests of the gardens themselves are by any means confined to taxonomy (see page 860). Many are also now involved in molecular research, much of it addressing spheres that are neglected elsewhere because they lack commercial significance. The agricultural or pharmacological potential of plants naturally dominates the research agendas of plant scientists in industrial companies and, increasingly, in university botany departments as well. That leaves the botanical gardens to study the astonishing diversity of plant species, their relations to each other and their evolutionary origins. That's a massive research agenda for which little public financial support is forthcoming.

Nonetheless, gardens have got on with their work in their usual unassuming way. Many of their botanists are able to follow their own interests more closely than their grant-dependent peers at universities. Intellectually, that's often a good thing, but politically it has crept into the gardens' institutional culture and has sometimes prevented them from engaging as fully as they might with the outside world. They often enjoy good relations with universities and the public, but there is room for even more active participation in debates on topics such as climate change, deforestation and urbanization.

Some ambitious and outward-looking projects are going forward at the larger gardens. The Royal Botanic Garden Edinburgh is planning

a £14-million (\$24-million) education centre, for example. And at the New York Botanical Garden next month, a \$23-million research centre will open whose work will include the investigation of the function of plant genes. Some academic botanists criticize this new direction as duplicative and a distraction from the garden's core work. But the genetic investigations pursued at botanical gardens are unlikely to overlap with university studies. Amy Litt, head of genomics at the centre, will investigate such diverse questions as the origins of seeds, the differences between hard and soft fruit, and the genes that determine the shapes of flowers, in part by sequencing genomes of plants such as the snapdragon.

The centre is part of the New York Plant Genomics Consortium, which includes the Cold Spring Harbor Laboratory, New York University and the American Museum of Natural History. The collaboration marks an effort by the garden to look beyond its own walls; it should be watched closely by those botanical gardeners who worry about possible marginalization within the wider research community. Some isolation has been self-imposed: there's a kernel of truth in the stereotype of the discipline as comprising a collection of reclusive eccentrics spending a lifetime on their particular interests.

As well as looking outside, botanical gardens need to work more closely with each other to

pursue the systematic collection of information about plants. Many gardens are loosely linked through an umbrella group, the Botanic Gardens Conservation International, based at Kew in London.

But this group focuses on conservation, not, for example, on the construction of databases. Several major gardens would each like their own systems to form the backbone of a global database. Closer cooperation is needed if knowledge of plants is to be integrated in a form that can be accessed by all botanists. It is past time for the largest gardens — at New York, Kew and St Louis — to work directly with each other to obtain financial backing for such a project. ■

"Botanical gardens need to work more closely with each other to pursue the systematic collection of information about plants."

More than the money

Technology-transfer offices are learning from their mistakes. So should the academics that they serve.

Academic researchers and the technology-transfer offices at their universities have had a prickly relationship since the latter first won a foothold on campuses two decades ago. Scientists sometimes complain that these offices are unresponsive to immediate demands, or that their generalist staff lack knowledge

in specific scientific or technical fields. Once they do start working together, researchers too often view the hapless technology-transfer officer as a potential obstacle to the dream deal they had been plotting with industrial partners or financiers outside the university. Yet university technology-transfer offices have come a long way. It is time that truculent researchers recognized their worth and engaged with them constructively, in the common interest of the university and its surrounding community.

Staff in these offices have, over time, built up valuable expertise in helping to negotiate deals with outside parties. Although people sometimes assume that the offices are just there to earn cash for the



university through royalty arrangements, the thinking of university administrators has moved on. It is now widely accepted that, aside from the occasional jackpot of the sort enjoyed by Columbia University in New York (whose 'Axel patents' for gene insertion have earned it more than \$300 million), technology-transfer offices are unlikely to generate large income streams. Instead, their principal role is to develop universities' ties with business in ways that should benefit students, staff and the surrounding community.

At last month's annual meeting of the Association of University Technology Managers, which drew some 2,000 technology-transfer officials from around the world to Orlando, Florida, the association's leadership declared: "It's not about the money." The meeting's busiest sessions were about the money, of course. But the point still stands: the remit of technology managers has grown far wider than just the collection of royalty payments.

For their part, technology managers find some researchers to be rather naive in their expectations regarding interactions with industry. Academics are sometimes slow to acknowledge potential pitfalls, such as pending ownership disputes over intellectual property. They can be too ready to sign over their future ideas in so-called 'honeymoon' deals, where a tightly controlled arrangement for the technology in hand would make more sense. More crucially, as technology managers see it, researchers take too sanguine a view of their would-be industrial partners, or their new-found venture-capital backers, on the other side of the table.

Like the technology managers, academic researchers are still on a

learning curve. But they know much more these days about business deals, and are aware of the broader role that scientific research plays in economic development.

The interests of the technology-transfer offices are broadening out. For example, Cambridge Enterprise has some 20 staff and expertise that reaches beyond patenting and licensing agreements to the distribution of seed capital to promising young firms (see page 867). Research universities all over the world are looking for the latter: seed money from private sources is hard to find, and various mechanisms are under investigation to keep it flowing (see *Nature* 440, 738–739; 2006). Even in technology hotspots such as Silicon Valley, a slowing flow of venture capital for early-stage company development makes it a topic for universities to address themselves.

Ideally, technology-transfer offices should be a trusted resource for university scientists, working to protect their interests and establishing the right kind of relationships with commercial partners. Some academics can do that for themselves, but most need professional assistance.

Politicians and industrial managers increasingly view the research university as an essential source of the innovative ideas that drive modern economies. University technology managers and academics should work together to make the most of their strong position. ■

"Ideally, technology-transfer offices should be a trusted resource for university scientists."

Neglected neighbour

Venus Express will go some way towards correcting a strange disparity.

Our two neighbouring planets form a stark contrast: one dedicated to love, a diamond in the evening, a pearl in the morning; the other dimmer, ruddier and tainted with bellicosity. You might think that each would get its fair share of our world's attention, but recently this has not been so.

The European Space Agency's Venus Express was due to enter orbit around Venus on 11 April, becoming its first Earthly visitor since Magellan, a NASA orbiter launched in 1989. In that interval, 15 separate spacecraft have been sent to probe the mysteries of Mars.

Yet Venus has plenty of mysteries of its own. Its high-level winds tear around at a speed no one can understand, and its surface bears the scars of an extraordinary cataclysm that overwrote most of its previous history less than a billion years ago. Scientists have developed little understanding of how this planet, so similar in size to Earth and yet so different from it, actually works.

In terms of maintaining Earthly interest, though, veiled Venus may have been mysterious for too long. The fact that the surface of Mars can be seen from Earth made it an appealing locale for fiction and fantasy. The astronomer Percival Lowell's vision of Mars as a desert planet, once Earth-like but now dying, gripped the twentieth-century imagination. As a distant desert, perhaps replete with strange and ancient inhabitants, Mars seemed to offer something akin to a

new frontier. So while Venus stayed a planet, Mars became a world.

Closer examination of Mars has shown that aspects of its geological history can be made to fit into this story, offering evidence compatible with a wetter and more habitable past that could also be taken as an invitation to an inhabited future. A mixture of science and story-telling thus made Mars the planet most studied for traces of life, and most speculated on as a destination for human explorers.

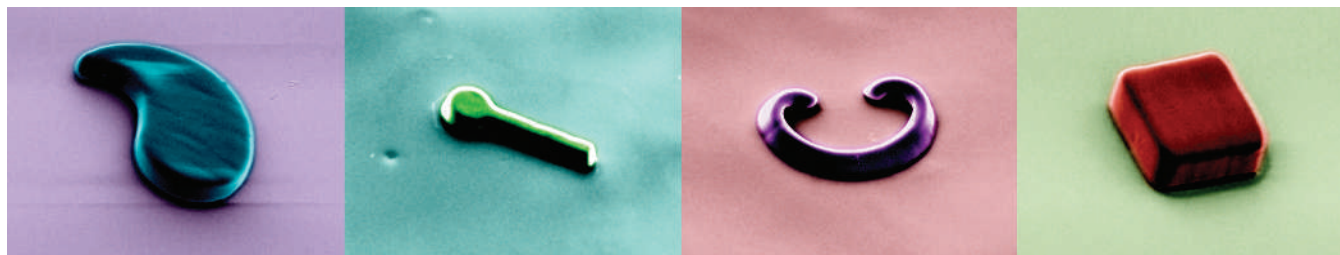
Venus fans point out that it, too, has possibilities along these lines. Like Mars, it must have enjoyed a less inclement youth, with oceans in which early life may have bathed. Unfortunately, Venus has erased almost all obvious signs of its earlier state. Even so, there are no rock-solid grounds for excluding the possibility that remnants of a venusian biosphere persist in its clouds to this day.

Nor is Venus an impossible candidate for human exploration. Its surface presents conditions more akin to the inside of an industrial chemical plant than to those encountered on a geological field trip. But journeys through its cloudscapes in dirigibles might still conceivably be less problematic than establishing a base on the cold, toxic, dust-devilled surface of Mars.

Without the narrative background into which our hopes for Mars are embedded, ideas of visiting Venus are unlikely to gain much traction. It is likely that Venus Express will yield gigabytes of fascinating data, greatly enriching knowledge of our nearest planetary neighbour. But if the mission is to usher in an age of Venus exploration, it will need to send back a compelling story too. ■

"Like Mars, Venus must have enjoyed a less inclement youth, with oceans in which early life may have bathed."

RESEARCH HIGHLIGHTS



Particles shape up

Nature Mater. doi:10.1038/nmat1617 (2006)

Swarms of polymer particles with curious shapes (pictured) can be

easily synthesized using a new continuous-flow technique from researchers at the Massachusetts Institute of Technology. Such particles may be useful building blocks for photonic crystals and biomaterials.

Patrick Doyle and his colleagues created the sweet-shaped objects, just a few micrometres in size, using materials that can be polymerized by ultraviolet light.

The material was pumped through the channels of a

microfluidic device while pulses of ultraviolet light were shone on to it through a mask defining the desired particle cross-section. This produced flat-sided particles, carried by the flow, at rates of up to 400,000 per hour.

GENETIC ENGINEERING

Blind mice

Neuron **50**, 23–33 (2006)

One strategy for restoring vision lost to retinal degeneration is to turn surviving nerve cells into light sensors. Could a light-activated protein from green algae do the trick?

Zhuo-Hua Pan at the Wayne State University School of Medicine in Detroit, Michigan, and his colleagues used viral gene transfer to insert the protein channelopsin-2 into the retinal ganglion cells of rodents. The protein formed light-gated channels, leading to light-evoked electrical responses in the neurons that are normally insensitive to light.

The team also injected the viral vector into the eyes of genetically blind mice. They recorded light-induced responses in the visual cortex as well as the retina, suggesting that the signals can be transmitted to the rest of the brain. But further work is needed to determine whether this constitutes 'vision'.

BIOCHEMISTRY

Pass it on

Chem. Biol. **13**, 329–338 (2006)

We already know that one kind of nucleic acid, such as DNA or RNA, can transfer nucleotide sequence information to another. Now researchers in the United States say that biochemical function can also be passed on.

Gerald Joyce and his team at the Scripps Research Institute in La Jolla, California, studied a particular ribozyme, an RNA molecule that contains genetic information and acts as an enzyme. They made a copy of the ribozyme that had the same sequence of nucleotides, but was instead made of DNA. The resulting enzyme did not function, but

could be evolved to do so with a few mutations. Something analogous to this process might explain how early life made the transition from a postulated predecessor of RNA to form the 'RNA world'.

ASTRONOMY

Heavenly alchemy

Astrophys. J. **641**, 1–20 (2006)

The Universe's first stars were much smaller than was once thought, according to a mathematical model of stellar chemistry.

Previous simulations of galactic evolution proposed that the first stars (called Population III) were more than 100 times as massive as the Sun. But such stars would live fast and die violently, producing none of the heavy elements such as barium that astronomers find inside the Population II stars that formed from their ashes.

Jason Tumlinson at the Yale Center for Astronomy and Astrophysics in New Haven, Connecticut, shows that Population III stars averaging between 10 and 40 solar masses would produce the observed chemical make-up of the long-lived second-generation stars, which are still around today.

MICROBIOLOGY

No sense of direction

Mol. Microbiol. **60**, 417–426 (2006)

Vibrio cholerae cells from human stools have previously been shown to be more virulent than their laboratory-grown siblings. Now Andrew Camilli of the Tufts University School of Medicine in Boston, Massachusetts, and his colleagues have traced the phenomenon to the genetic modification of a single trait.

Camilli's group had shown before that expression of certain genes was reduced in

stool-isolated *V. cholerae*. In this study, the team finds that the stool bacteria are unable to direct their swimming towards nutritional cues — a process known as chemotaxis. They linked this to reduced expression of *cheW-1*. It remains unclear, however, why impaired chemotaxis reduces the number of *V. cholerae* cells required for a successful infection.

EVOLUTION

March of the ants

Science **312**, 101–104 (2006)

Today's taxonomic ant groups arose earlier than previously thought, more than 140 million years ago, according to the most comprehensive census yet of their evolutionary family tree. But they may not have diverged into the 11,800 species that now exist until much later — perhaps at the same time as flowering plants diversified and increased the ants' range of lifestyle options.

Researchers led by Corrie Moreau of Harvard University in Cambridge,



S. B. ARCHIBALD/HARVARD UNIV.

Massachusetts, compared corresponding sets of DNA sequences from a huge range of ant species. By evaluating the divergence between the sequences, they deduced how long ago the different groups emerged — and then went back to the fossil record (pictured) to check.

PHYSIOLOGY

What it takes to hibernate

Cell **125**, 161–172 (2006)

Chipmunks have helped researchers to crack a tough nut — finding the proteins that control hibernation.

Noriaki Kondo of the Mitsubishi Kagaku Institute of Life Sciences in Tokyo, Japan, and his colleagues watched more than 100 male chipmunks (*Tamias sibiricus*) go through their annual hibernation cycles, the oldest living for 11 years. Their patience paid off.

First, by keeping the chipmunks at a constant 5 °C, the researchers showed that hibernation occurred independent of temperature. Second, they discovered that levels of hibernation-specific protein complex (HPC), which is known to drop off in the blood at the beginning of hibernation, rose in the brain during the same time. Blocking HPC with antibodies either stopped or shortened hibernation.

DEVELOPMENTAL BIOLOGY

Baby's first steps

Dev. Cell **10**, 451–459 (2006)

A molecular signal that distinguishes the mouse embryo's future head from tail appears earlier than had been thought, and without directions from the uterus.

The first clear sign of the body's axes was thought to emerge after five days of development and after implantation. Hiroshi Hamada of the Osaka University, Japan, and his colleagues show that a gene called *Lefty1* is switched on in the future head region of embryos at around four days — and that this also occurs when embryos are grown in culture, without the guidance of the uterine wall.

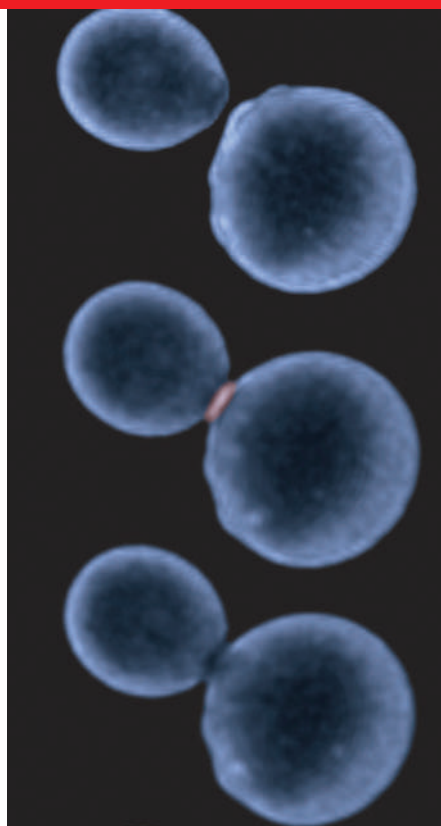
All mammalian embryos may use a similar mechanism to determine their body plan, the authors say.

CELL BIOLOGY

Cutting edge

Cell **125**, 85–98 (2006)

Yeast cells have a signalling system that ensures chromosomes are safely hauled to their correct destinations before allowing the final step of cell division to go ahead, researchers have found.



Yves Barral of the Swiss Federal Institute of Technology in Zurich and his colleagues studied abscission in budding yeast (pictured above), the process during which the membranes between two nascent daughter cells pinch off. They found that a signalling system, which they called NoCut, delays abscission until the replicated chromosomes have been pulled apart into the daughter cells and are well clear of the closing membrane.

The system helps to stop chromosomes being broken during division. It could also be at work in mammalian cells, where it would be an important guard against the development of cancer.

SEMICONDUCTORS

Full of holes?

Phys. Rev. Lett. **96**, 125505 (2006)

After years of effort, a delinquent semiconductor has been tamed. Researchers in the United States report evidence for p-type doping of indium nitride (InN).

P-type doping introduces positively charged conductors called 'holes' into a semiconductor. This is necessary for building devices such as solar cells, long-wavelength lasers and fast transistors, for which InN holds promise. But efforts to make p-type InN have been hampered by problems with measuring the result: surface electrons make the definitive test impossible to carry out.

Now researchers at the Lawrence Berkeley National Laboratory, California, and Cornell University in Ithaca, New York, have done a battery of experiments that suggest adding magnesium to the material does the job.

JOURNAL CLUB

Uta Passow

Alfred Wegener Institute,
Bremerhaven, Germany

There's more to marine snow than meets the eye, says a biological oceanographer.

I will never forget my first dive in the open ocean. Weightlessly suspended in the blue waters of the Pacific, I was surrounded by a quiet, gentle drizzle of falling marine snow. The beauty of it almost took my breath away.

Marine snow (the subject I study) consists of composite particles, more than half a millimetre in diameter, made of cells, shreds of plankton feeding webs, faeces and minerals. It represents a critical link in the global carbon cycle: the sinking snow transports carbon from the upper ocean, and thus the atmosphere, to the ocean depths.

In calculating the snow's contribution to the carbon budget, scientists had assumed that the carbon flux depended only on the particles' solid content. But new evidence may force us to revise this view, leading to the conclusion that marine snow transports more carbon than previously believed.

Microbes degrade the solid matter in the snow as it descends, releasing carbon into the water. It was thought that this dissolved organic carbon would be left behind, so reducing the carbon flux with depth.

However, recent measurements have found high dissolved-organic-carbon levels in the funnel-shaped sediment traps used to collect marine snow, suggesting that some of the dissolved carbon can hitch a ride downwards (A. N. Antia *Biogeosciences* **2**, 189–204; 2005).

Independent work supports this idea, demonstrating that marine-snow aggregates somehow retain solutes within their pores, despite being 99% holes (S. A. Goldthwait *et al. Mar. Ecol. Prog. Ser.* **305**, 59–65; 2005). Clearly, the mysteries of the tiny islands of activity called marine snow are far from solved.

NEWS

Bird-flu experts question advice on eating poultry

Can people catch the H5N1 avian flu virus from eating infected poultry? Colin Blake-more, chief executive of the UK Medical Research Council, says the public need not worry. "There is no evidence of transmission to people by eating cooked eggs or chicken," he said on BBC radio last week, adding that the only food risk he could see was from "drinking swans' blood".

Blakemore's sound bite came a day after Britain's first case of H5N1 in a wild bird was confirmed — a dead swan found floating in a harbour in Cellardyke, Scotland. And it echoes a slew of recent reassurances by governments worldwide and by the World Health Organization (WHO), all conscious of damaging public confidence and the poultry industry.

But many flu scientists are concerned that, although the risks are low compared with those associated with contact with diseased birds, there is not enough evidence to say that the virus cannot be transmitted by eating infected poultry. "Oral transmission is an open question," says Masato Tashiro, a virologist at the National Institute of Infectious Diseases in Tokyo. "Direct evidence of oral infection is lacking, but so too is proof against."

On 23 March, the European Food Safety

Authority (EFSA) published a prominent scientific risk assessment (*EFSA J.* 74, 1–29; 2006). Its advice is that poultry products are safe to eat and have "not been implicated in the transmission of the H5N1 avian influenza virus to humans".

H5N1 is present in the meat and eggs of infected birds, and animals have become infected by eating diseased birds. But the EFSA plays down this route in humans, arguing that "humans who have acquired the infection have been in direct contact with infected live or dead birds".

That overstates the case, says Jody Lanard, a physician and risk-communication consultant based in Princeton, New Jersey, who has recently advised the WHO about pandemic communication. She points out that the report itself acknowledges elsewhere that in many instances there is not enough epidemiological evidence to identify the source of infection, and that poor preparation and cooking of food cannot be excluded as the cause. "Such cases could equally well indicate a likely gastrointestinal portal of entry," agrees Menno de Jong, a virologist at the Hospital for Tropical Diseases in Ho Chi Minh City, Vietnam.

The report also dismisses the idea that H5N1

"Direct evidence of oral infection is lacking, but so too is proof against."



Don't panic: official advice is that despite bird-flu fears, chickens are safe to eat.

F. VILLA/AFR/GETTY

can enter the body via the human gut, concluding that there is "no proof that virus replicates in the human intestine". Although it mentions the presence of diarrhoea in infected humans, together with the detection of viral RNA in intestines and the virus in rectal swabs, it says these "do not allow one to conclude that the GI tract is a portal of entry or a target organ".

De Jong, who has treated many of the cited diarrhoea cases, says the report's authors are

Patients warned about unproven spinal surgery

Pressure is mounting on a Beijing neurosurgeon to prove that his popular treatment for spinal-cord injury works. In an article published last month, a group of spinal experts concludes that the treatment, which involves implanting fetal cells into the spine to promote nerve-cell regeneration, has significant side effects and does not provide any benefit.

Hongyun Huang's technique is based on the theory that olfactory ensheathing cells (OECs), which normally help to link nerve cells in the nose and the brain, can help regenerate nerve cells at the site of an injury. Since 2001, Huang, who

works at Chaoyang Hospital, has treated around 600 patients with tissue from aborted fetuses that he says contain OECs (see *Nature* 437, 810–811; 2005).

Three spinal-cord experts have now published a critique of his methods (B. H. Dobkin, A. Curt and J. Guest *Neurorehabil. Neural Repair* 20, 5–13; 2006); the researchers followed seven of Huang's patients before and after treatment, reviewed his publications and visited his lab.

They say Huang's surgical techniques are good. But despite Huang reporting at a February 2004 meeting that there had been no adverse effects of more than 500

implantations, the researchers found that five of the seven patients experienced side effects including meningitis. They also question whether the cells used by Huang are OECs. "We don't know what those cells are but they are not pure OECs," says first author Bruce Dobkin, a neurorehabilitation specialist at the University of California, Los Angeles.

The most damning claim in the critique is that none of the patients showed any improvement after treatment. For example co-author Armin Curt, a neurologist at the International Collaboration on Repair Discoveries (ICORD) at the University of British Columbia,

Canada, measured muscle responses to nerve signals, and found no change in the three patients he studied before and after the procedure. "There is a good chance patients are just wasting money and expectations," says Curt.

Huang vigorously defends his technique and calls the report a "pack of lies", pointing out that the third author, James Guest of the Miami Project to Cure Paralysis, has previously reported observations that support his procedure (J. Guest, L. P. Herrera and T. Qian *Spinal Cord* 44, 135–142; 2006). Several patients contacted by *Nature* claim that they have

**BIRD FLU ON THE MOVE**

Visit our website to find a timeline of the H5N1 virus's spread from Asia to Europe.

www.nature.com/news



“formally right” to say there is no proof that the virus replicates in the intestine. But there is no proof that it doesn’t either, he says, noting that some of the diarrhoea cases had no respiratory symptoms.

“Available evidence suggests that the gastrointestinal tract in humans is a portal of entry for H5N1,” agrees Albert Osterhaus, a virologist at the Erasmus Medical Center in Rotterdam. He carried out a recent study in which cats became infected with H5N1 after being fed infected chickens (T. Kuiken *et al. Nature* 440, 741–742; 2006) — evidence also dismissed

by the EFSA report as “unproven”.

Of course, to pose a risk the virus must enter the human food chain. The EFSA and other authorities point out that this is unlikely, at least in industrialized countries. But some scientists, including Osterhaus, say it cannot be excluded — for example, if the virus enters poultry a few days before clinical signs appear.

The final argument of the EFSA and the WHO is that even if the virus did enter the food chain, it would be killed by cooking or pasteurization, in the same way as bacterial pathogens such as *Salmonella*. But Lanard

points to the countless infections with *Salmonella* worldwide, and complains that most risk assessments fail to acknowledge that in reality few people follow guidelines for the safe handling and cooking of poultry. These involve cooking chicken right through to 70 °C and eggs until they are hard, using separate knives and chopping boards for raw and cooked foods, and hand-washing between operations.

An EFSA spokesperson says the agency stands by the report’s conclusions. Several scientists are also convinced that avian flu carries no food risks. “Avian influenza has never been and should never have been seen as a food safety issue,” says Les Sims, a consultant for the UN’s Food and Agriculture Organization (FAO). Bird-flu concerns over food, he says, “have a devastating impact on the livelihood of millions of farmers globally and demonstrate that risk communication on this has been a total failure”.

But Lanard maintains that to say bird flu is not a food issue is an “overstatement”. She says that such advice shows that little has been learnt about risk communication since the British agriculture minister publicly fed his young daughter a hamburger at the height of the crisis over bovine spongiform encephalitis. A 2005 European Commission poll showed that almost half of European citizens believe authorities favour economic interests over consumer health, she points out. “These over-reassuring statements discount the future — they are set up for public distrust,” she says. “Although there is no direct evidence that transmission can occur through poorly cooked infected poultry, all animal evidence to date unfortunately suggests that this is possible.”

Declan Butler

experienced benefits. A nurse practitioner in California says her son showed various improvements, such as reduced sweating, muscle spasm and pain, and a better sense of balance since he had the surgery in November 2004, and accuses Dobkin of focusing on particularly difficult cases.

But Dobkin says such testimony can be misleading, because patients are so desperate to believe that they have improved after forking out US\$20,000 (or \$3,700 for Chinese patients) and undergoing a risky procedure. “They forget what they were like before,” he says. The surgery itself could also lead to short-lived improvements by relieving pressure in the area, adds John Steeves, a spinal-cord injury

specialist at ICORD. “You really need long-term follow-up,” he says. Steeves, who has visited Huang’s lab, says Huang has consistently ignored advice on how to use blind assessment, randomized controls and long-term observation.

Other spinal-cord treatments offered in countries such as Brazil and Portugal are also gaining popularity despite the lack of such data. This has prompted a group of international researchers to draft guidelines to help patients and clinicians evaluate treatments. Sponsored by the International Campaign for Cures of Spinal Cord Injury Paralysis, the guidelines

“There is a good chance patients are just wasting money and expectations.”

include recommendations and hard data to help distinguish the effects of a given treatment from other factors, says contributor Mark Tuszynski, a neuroscientist at the University of California, San Diego. For example, 40% of individuals show some spontaneous recovery after acute spinal-cord injury, so anecdotal stories of improvement don’t necessarily mean a treatment works. The guidelines will be submitted to an academic journal this summer, and a simplified version translated into many languages.

But getting the word out may not be easy. Paul Lu is a postdoctoral student in Tuszynski’s lab who

entered neurology after he was paralysed from the waist down in a car accident. A Chinese native, Lu has been trying to warn patients in China about Huang’s procedures, but says his translation of the *Neurorehabilitation and Neural Repair* article was rejected by several journals and science-based newspapers. “Chinese journals like to claim that China is leading the world in cell transplants,” says Lu. “They’d lose face if they print this.”

Last week, however, the *Chinese Journal of Spine and Spinal Cord* agreed to publish the translation in June. Readers will have to balance such reports with Huang’s confidence, buoyed by patient testimony, in his therapy. **David Cynoski**

ON THE RECORD

“We are Earth scientists. We are not part of a vast conspiracy to perpetrate a hoax, nor are we crowd-following herd animals.”

Oceanographer David Archer of the University of Chicago refutes conservative US columnist Robert Novak's claim that scientists are hyping global warming.

“To sell the Yucca Mountain project to our children through the use of a cartoon character is an irresponsible and desperate act.”

Congressman Jon Porter (Republican, Nevada) calls on the government to dismiss 'Yucca Mountain Johnny', a cartoon mascot promoting the plan for a nuclear-waste dump in Nevada.

Sources: RealClimate.org, Las Vegas Review-Journal

SCORECARD



Artificial sweetener
Researchers at the US National Cancer Institute

find that aspartame, a known carcinogen in rats, does not increase the risk of cancer in people.



Christ's buoyancy
An oceanographer at the University of Florida

suggests that Jesus managed to walk on water because the Sea of Galilee was frozen.

NUMBER CRUNCH

Health workers are one of the most important factors affecting infant mortality, according to figures from the World Health Organization.

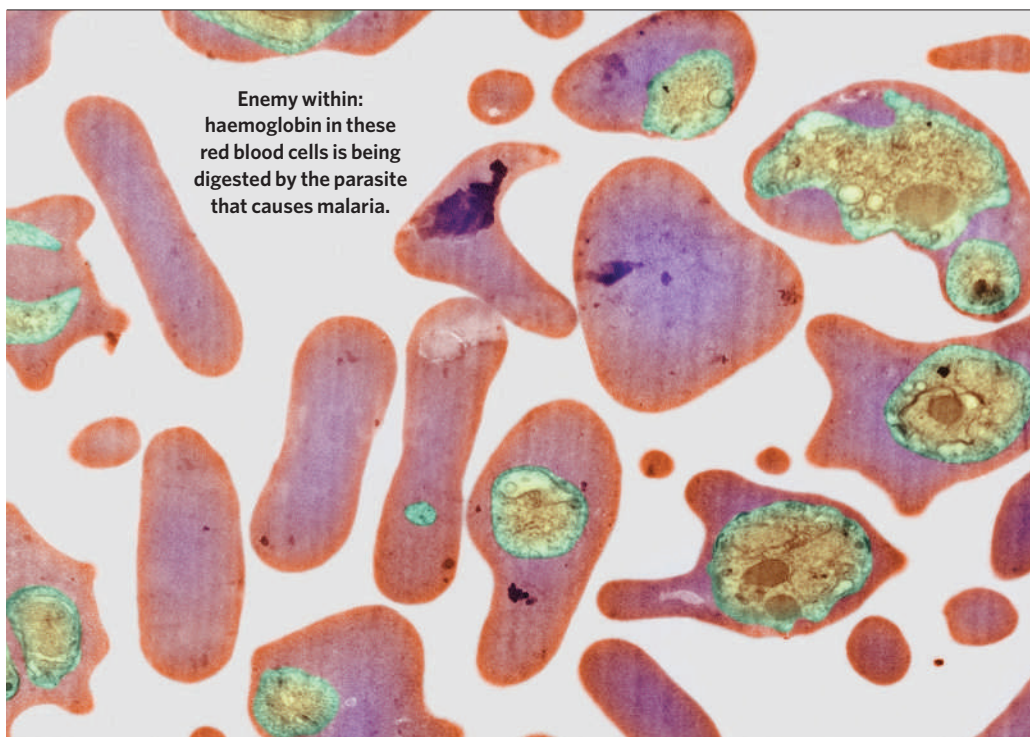
0.28 is the number of physicians per 1,000 citizens in Nigeria.

19.7% is the chance that a Nigerian child will die before the age of five.

2.3 is the number of physicians per 1,000 citizens in Britain.

0.6% is the chance that a British child will die before the age of five.

Source: WHO The World Health Report 2006.



Malaria breakthrough raises spectre of drug resistance

The 'miracle' malaria drug artemisinin is a step closer to being produced plentifully and cheaply. Synthetic chemists have put plant genes into yeast to make it churn out large amounts of the precursor artemisinic acid. The discovery brings hope to areas such as sub-Saharan Africa, where those who need the drug most can ill afford it.

Researchers have praised the work and are excited that it may soon be possible to get artemisinin to the 300 million to 500 million people infected with malaria each year. But many are also concerned that this will trigger the emergence of resistance to the drug, thus destroying our most effective weapon against the disease.

Artemisinin is extracted from the leaves of *Artemisia annua*, or sweet wormwood, and has been used for more than 2,000 years by the Chinese as a herbal medicine called *qinghaosu*. The parasite that causes malaria has become at least partly resistant to every other treatment tried so far. Artemisinin is still effective, but it is costly and scarce.

"This drug is such an important thing for malaria," says David Warhurst, an expert in

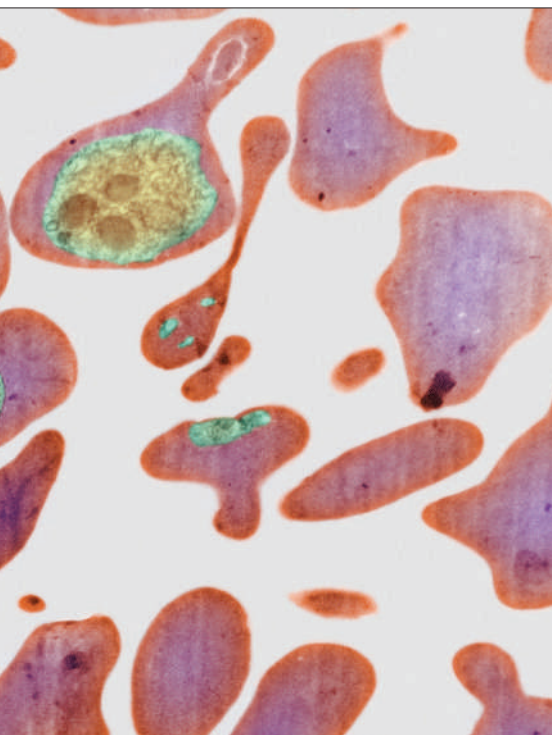
protozoan chemotherapy at the London School of Hygiene and Tropical Medicine. "It is the basis of all the new treatments that are going ahead."

Jay Keasling at the University of California, Berkeley, and his colleagues tweaked a pathway and used three plant genes to persuade yeast (*Saccharomyces cerevisiae*) to produce and secrete large amounts of artemisinic acid, which is just a few chemical steps away from artemisinin. The researchers, whose paper starts on page 940 of this issue, hope that once

the process is scaled up it will allow artemisinin to be made industrially. A course of artemisinin currently costs US\$2.40; cutting the cost to 10% of that should make it affordable for most sufferers.

Work towards industrial production has already been started by the non-profit pharmaceutical company Institute for OneWorld Health, based in San Francisco, California, in partnership with Amyris Biotechnologies and with the help of \$42.6 million from the Bill and Melinda Gates Foundation. "We're focusing on producing a known pharmaceutical so that it reaches the people who need it most," says Jack Newman, co-

"Artemisinin is the basis of all the new treatments that are going ahead."



LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE/SPL

founder of Amyris, based in Emeryville, California. But he is hopeful that the work will lead to industrial production of other plant-based drugs in a similar way.

The prospect of plentiful artemisinin is encouraging, but if the parasite becomes resistant, increased drug production will be worthless. "The loss of artemisinin could spell disaster for malaria sufferers," warns Chris Hentschel, head of the non-profit Medicines for Malaria Venture, based in Geneva.

There is no consensus on how likely resistance is, but some think the risk is high. Artemisinin works by disabling a calcium pump in the malaria parasite, and last year researchers reported that the mutation of a single amino acid was sufficient to confer resistance (A.-C. Uhleman *et al. Nature Struct. Mol. Biol.* **12**, 628–629; 2005). When another team gave low doses of artemisinin to parasites taken from patients in French Guiana, some mutated, becoming highly resistant to the drug (R. Jambou *et al. Lancet* **366**, 1960–1963; 2005).

The main way to stop resistance arising is to always give the drug in combination with others. In January, the World Health Organization (WHO) made a plea to pharmaceutical companies to end the marketing and sale of single-drug artemisinin medicines. But as other malaria drugs grow increasingly ineffective, many feel that resistance to artemisinin is inevitable.

"We hope it won't happen," says Warhurst. "But looking for new drugs is important." ■

Narelle Towie

Two telescopes join hunt for ET

The search for extraterrestrial intelligence (SETI) will ramp up in coming months as two dedicated facilities come online — one to look, the other to listen.

A team led by physicist Paul Horowitz of Harvard University will begin scanning the skies this week for flashes of light from alien civilizations. Most SETI searches have been at radio wavelengths, but theorists surmise that extraterrestrials might also shine laser beacons visible from a few thousand light years away.

This will be the first optical SETI project to scan the entire sky, or at least all that can be seen from the Oak Ridge Observatory in Harvard, Massachusetts, where a 180-centimetre telescope has been installed. The \$50,000 instrument was paid for by the Planetary Society, a grassroots group of space enthusiasts, and will record flashes briefer than a nanosecond. No known natural process causes such flashes. The all-sky search requires 200 nights of clear viewing, and is expected to take several years.

Meanwhile, at the Hat Creek Radio Observatory in northern California, the first ten dishes of the privately funded Allen Telescope Array are due to be demonstrated later this month, says Peter Backus, observing programmes manager at the SETI Institute in Mountain View, California. With money from Microsoft alumni Paul Allen and Nathan Myhrvold, the institute, working with Berkeley, is building an array of 350 six-metre radio dishes dedicated to SETI. The entire array will eavesdrop on nearly a million stars for hints of intelligence.

Managers hope to have 42 antennas working by July. The first scan will be of a narrow swathe of the Milky Way's centre. ■
Tony Reichhardt



I. GARY

See the future: how the Allen array might look.

**IS IT A DREAM?**

People who have near-death experiences are more likely to find REM sleep intruding on reality.

www.nature.com/news

Can super-antibody drugs be tamed?

As it becomes clear that the London clinical trial disaster was indeed the fault of the drug itself, **Michael Hopkin** looks at what went wrong, and whether there is any future for 'superagonist' antibody therapies.

There was no warning from animal tests, but last month the experimental antibody drug TGN1412 put six British men in intensive care. The resulting investigation ruled out any failure of experimental and regulatory procedures — a relief for those involved, but a damaging blow for the field. Immunologists are now left asking what went so badly wrong in the trial, and whether the fearsome potency of 'superagonist' antibody therapies will outweigh their promise.

The six volunteers at London's Northwick Park Hospital were probably struck by a huge immune reaction called a cytokine storm — a flood of inflammatory molecules released by cells called helper T cells, which shut down their organs in hours (see *Nature* **440**, 338–339; 2006). The UK Medical and Healthcare products Regulatory Agency, which approved the trial, announced last week that it has found no evidence of contamination in the treatments, which means the devastating effects were almost certainly caused by TGN1412 itself.

With hindsight, it might be no surprise that the compound, dubbed a 'superagonist' antibody by its creators, could run amok in the immune system. Around 20 antibody therapies are currently approved or nearing approval, most of which mimic natural human antibodies against specific viruses or cancer-cell types. But TGN1412 is different. It was designed to circumvent the usual checks and balances that prevent T cells from overreacting in the course of their normal duties.

Usually, T cells respond to specific enemies, or antigens, recognized by the immune system. After the antigen is first encountered, an 'antigen-presenting cell' binds to a T cell at two sites on its surface: the T-cell



Gathering storm? Antibody therapies aimed at large swathes of the immune system might be too hot to handle.

"With hindsight, it might be no surprise that the compound could run amok in the immune system"

receptor (TCR) and the CD28 receptor. TCR binding is specific to the antigen in question, whereas CD28 binding acts more like a 'go' switch for the T cell; both are normally required (see graphic, overleaf). But the makers of TGN1412 found a way to switch on the CD28 green light without TCR binding, activating T cells across the board.

T cells fall into three broad categories: killers, which destroy specific target cells; helpers, which boost other parts of the immune system and are the most likely source of a cytokine storm; and regulatory T cells, which suppress other

elements of the immune system. Roughly half of killer T cells, and virtually all helper and regulatory T cells, express the CD28 receptor.

When Thomas Hünig, an immunologist at the University of Würzburg and researchers at TeGenero, the spin-off company he co-founded to develop TGN1412, began testing the antibody on animals, they found that the drug seemed only to activate regulatory T cells (Beyersdorf, N. *et al. J. Exp. Med.* **202**, 445–455; 2005). This made it an enticing candidate for treating autoimmune diseases such as rheumatoid arthritis and type-1 diabetes. The researchers hoped that TGN1412 could make immunosuppressive cells soothe sites of overinflammation, while the rest of the immune system carried on as usual.

It is still not clear why, when TGN1412 affected only regulatory T cells in animals, the same almost certainly did not occur in the human trial. It is likely that in humans the super-antibody activated helper T cells en masse, triggering a cytokine storm.

"We were shocked and surprised to see what happened in humans," Hünig told *Nature*. In preclinical trials, monkeys got a

dose 500 times that given to the human volunteers, and the monkey CD28 receptor is identical to the human one, says Hünig. This means that the effects in the monkey trial should have been comparable.

One possible source of the difference between the animal and human trials is that the 'tail' of the antibody molecule at the opposite end from the CD28-binding site may not be the same in humans and monkeys. Antibody tails can undergo a process called crosslinking, which amplifies an immune response by recruiting more immune cells or antibodies. Therapeutic antibodies are modified so that they have the same overall structure as a generic human antibody, but doing

J. HOLMES/CELLTECH/SCIENCE PHOTO LIBRARY

this may also have prevented the full extent of TGN1412's activity from showing up in animal tests. Hünig admits that this could have happened.

Hünig's team should have known of previous research on this region, says Camilo Colaco of ImmunoBiology, an immunological research company in Cambridge, UK. Similar work has been done on CD3 receptors, which are crucial for the TCR component of T-cell activation, he says. Researchers were hoping that a superagonist antibody directed at this receptor could boost general T-cell activity, helping transplant patients restore their immune systems after taking immunosuppressive drugs.

Indiscriminate killers

But in tests of the antibody in mice, researchers found that uncontrolled cytokine release was a problem — albeit not a large one because the mice were already immunodepleted (L. Chate-noud *et al. Transplantation* **49**, 697–702; 1990). Tweaks to the antibody's tail to prevent crosslinking resulted in an antibody that did not have the same problem and could enter clinical trials (P. A. Carpenter *et al. Biol. Blood Marrow Transplant.* **11**, 465–471; 2005); the drug has now been approved under the name visilizumab. But using unmodified superagonists in a healthy volunteer with an intact immune system is playing with fire, says Colaco: "It would go bananas."

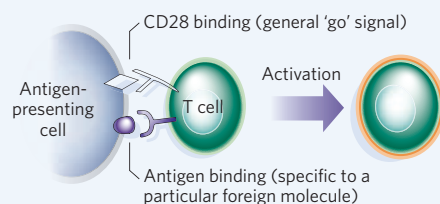
Another, more controversial, theory is that, in humans, CD28 receptors are found in the immune system beyond just T cells. Dorothy Lewis, an immunologist at Baylor College of Medicine in Houston, Texas, says that the receptor may be present on granulocytes, the most numerous white blood cells in the body, which have a range of cell-killing tricks. Alternatively, killer T cells might have been triggered to release a cell-destroying molecule called granzyme B. These processes could theoretically produce the effect seen in the TGN1412 trial, says Lewis. She also suggests that a cytokine storm might come on too slowly to account for the reaction seen in the trial volunteers, all of whom suffered severe effects within minutes and collapsed within 12 hours of being injected with TGN1412. "When you trigger T cells it takes a while to get up to speed," she says. "But granulocytes are good to go."

Hünig doesn't buy her theory, however. After the trial went wrong, he says he went over the literature to see if CD28 receptors might be more common than he had imagined. He says that only one group has ever claimed to have spotted them

"My first reaction was 'how did they get that through the regulatory body?'"

Normal T-cell activation

To switch on a T cell, two signals are required: a general activation trigger and a specific 'antigen' from a molecule recognized as being foreign.

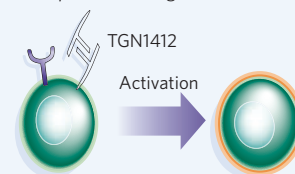


What happened in animal trials

In monkeys, TGN1412 activated only regulatory T cells, which damp other elements of the immune response. So the researchers hoped the drug would help treat autoimmune diseases by suppressing the inflammatory immune cells (red) that attack the patient's own tissue.

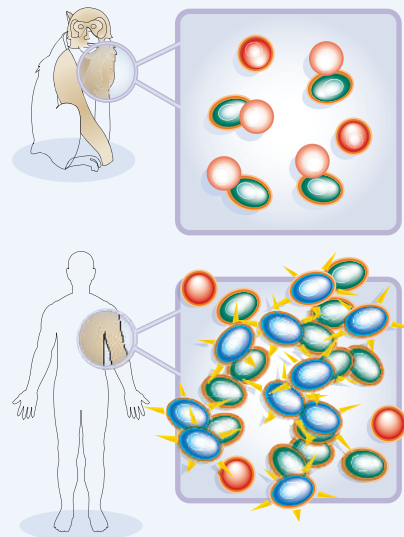
Super-antibody T-cell activation

This overrides both signals and potentially activates all T cells with a CD28 receptor, not just those specific for a particular antigen.



What probably happened in the London trial

As well as activating regulatory T cells, TGN1412 switched on helper T cells (blue), which produce chemical messengers called cytokines that boost other elements of the immune response. This mass activation of T cells caused a devastating 'cytokine storm' — a flood of inflammatory molecules that swept through the patients' bodies.



on granulocytes (K. Venuprasad *et al. Eur. J. Immunol.* **31**, 1536–1543; 2001), and he is confident that the idea can be ruled out. Hünig remains convinced that overactivation of helper T cells was to blame: "It was not too fast to be cytokines."

Proceed with caution

Some, such as Steve Anderton, an immunologist at the University of Edinburgh, UK, who is studying the use of antibodies to overcome cancers that evade the immune system, are concerned at the haste with which TGN1412 was tried out in humans — barely a year after the structure of the CD28 receptor was deduced (E. J. Evans *et al. Nature Immunol.* **6**, 271–279; 2005). "It is flabbergasting that this approach was tried," he says. "It takes away all the rules of the immune response." Jim Riley of the University of Pennsylvania in Philadelphia, who works on the CD28 pathway, feels similarly. "My first reaction was 'how did they get that through the regulatory body?'" he says.

A major question now is how the field can move forward. Instead of risking further work with superantibodies, some favour turning to other approaches. One option is to culture a patient's regulatory

T cells in the lab, then inject them into the bloodstream. Another is to target processes such as the binding of antigen-presenting cells to T cells using blocking antibodies. This method allows autoimmune disease to be treated without activating regulatory T cells, and has already been tested on a mouse model of multiple sclerosis (H. D. Lum *et al. Leukocyte Biol.*; in the press).

But others feel that superantibody therapy does have a future. The CD3 superantibody visilizumab suggests that such drugs can be safe. And Anderton says that the CD25 receptor, which is expressed mainly by regulatory T cells, and by other T cells only once they are activated, may be a safer target than CD28.

TeGenero's future is unclear — it has a staff of just 15 people, and TGN1412 was its only product. The company seems hopeful that the technology can be tamed: "It is too early to give definitive answers on how superagonistic antibodies must be developed in the future," TeGenero said in a statement to *Nature*. "Dangers can possibly be reduced by very careful assessment of pharmacological as well as safety characteristics."

But the idea of using superantibodies to activate the CD28 receptor is almost certainly dead. "In its current form I wouldn't want to have anything to do with it," says Riley. ■


**ARCTIC WATER FLOW
SPEEDING UP**

Measures of rain, snow and runoff in Siberia show hefty changes.

www.nature.com/news

Japan revises its military plans for space

TOKYO

Japan's ruling party is poised to revolutionize the country's space programme by allowing it to expand into military applications. In a country constitutionally wedded to peace, the move could cause public uproar. Scientists are also worried that they might lose out.

On 28 March, a committee of the Liberal Democratic Party proposed that Japan should revise a 1969 policy limiting the use of space to "peaceful" objectives. Under the current pacifist constitution, Japan has shied away from developing high-tech military satellites even for defence.

The committee suggests that Japan should lift its self-imposed limit on the imaging resolution of military satellites. It plans to submit a bill to the Diet next year to change the way space development is defined, which is likely to pass.

"Space development should be a source of national pride, but Japan doesn't have a diplomatic strategy to take advantage of it," says

Kenichi Kawamura, a spokesman for committee head Takeo Kawamura.

The government is concerned about possible attacks from terrorists and from North Korea, which fired ballistic missiles into Japanese water in 1998. Since then, Japan has launched two spy satellites, but their resolution is a quarter of that of other countries' reconnaissance satellites, and Japan has relied on the United States for high-resolution images and communications. The new law would enable Japan's defence agency to have spy satellites with a resolution that could detect a missile launch by another country.

The move is being welcomed by Japan's space industry, which hopes that new investment will make satellite manufacturers more competitive internationally. But there are worries abroad. "US government officials are likely to have great concern about Japan's development of its own

military satellites because they are nervous of the spread of understanding of technology," says Howard McCurdy, a space-policy expert at the American University in Washington DC.

Japanese researchers are also worried. They fear that although the Japan Aerospace Exploration Agency (JAXA) might enjoy a budget

boom in defence-related technologies, smaller-scale science missions could be cut back if government money was funnelled into the space industry. "Research areas that can be applied to defensive use would be in the spotlight,"

says Yasunori Matogawa, JAXA's associate executive director, adding that even areas where Japan has traditionally been strong, such as X-ray astronomy and infrared technologies, may be affected. "Those technologies look close to military ones but they are not the same." ■

Ichiko Fuyuno

"Space development should be a source of national pride, but Japan doesn't have a diplomatic strategy to take advantage of it."

'Nano' cleaning product recalled after health scare

A German company has mounted what is thought to be the first ever withdrawal of a nanotechnology-based product.

Late last month, more than 100 people reported headaches, insomnia and breathing problems after using two versions of a cleaning product known as Magic Nano. German discount store Penny Markt pulled the products from its shelves when it learned of the complaints.

At least six cases required hospitalization, according to Germany's Federal Institute for Risk Assessment in Berlin. The agency says it is investigating the cause of the illness, and it is unclear whether the 'nano' component of the product was involved.

The recall comes just weeks after a US survey showed a sharp rise in the number of consumer products advertised as nanotechnology (see *Nature* 440, 262; 2006).

Californian stem-cell centre gets stopgap funds

California's first training grants for stem-cell research is being funded with \$14 million in bond anticipation notes, which have been sold to six philanthropies, officials announced this week. These short-term bonds are intended to be paid off by a larger bond issue.

The California Institute for Regenerative Medicine will use the money as interim funding to cover training programmes identified in a review last year (see *Nature* 437, 800–801; 2005).

In 2004, Californian voters approved \$3 billion for stem-cell research, but the state has not been able to sell the bonds

because of lawsuits by religious groups and others that accuse the state of conflicts of interest and inadequate oversight. A court decision on the case is due soon.

If the institute wins, it hopes to begin selling the \$3 billion in bonds by spring 2007, repaying the bond anticipation notes now being purchased by the charities. Another \$36 million worth of bond anticipation notes are being sold to philanthropies for more interim research funding, officials say.

NASA reveals striking plan to find water on the Moon

NASA will smash a spent rocket stage into the Moon in 2008 to try to answer the long-standing question of whether water ice exists inside permanently shadowed lunar craters.

The crash will be monitored by the Lunar Crater Observation and Sensing Satellite (LCROSS), which was selected this week as a piggyback experiment on the Lunar Reconnaissance Orbiter (LRO). The orbiter is being sent to create detailed maps of the Moon ahead of the expected arrival of astronauts in 2018.

Having delivered LCROSS to the Moon, the 2,000-kilogram empty rocket stage will crash into Shackleton Crater near the lunar south pole. LCROSS will then examine material thrown up during the impact and fly through the plume, looking for spectroscopic evidence of water vapour or ice. The satellite itself will then crash into the lunar surface generating a second plume for further analysis. Telescopes on Earth and the LRO itself will also be focused on the impact site. NASA planners hope to find water that can be used as a resource for a future Moon base.



M. EVANS/AP

Successful: John McDarby has won substantial damages in his case against Merck.

Jury rules against Vioxx in heart-attack case

The 2004 heart attack of 77-year-old John McDarby, a retired insurance agent, was due partly to taking the painkiller Vioxx for four years, a New Jersey jury decided last week. The jury — from the home state of Vioxx's maker, Merck — awarded McDarby \$4.5 million in compensatory damages and was mulling punitive damages as *Nature* went to press.

The same jury found Vioxx not to blame for the heart attack of Thomas Cona, 60. Cona could provide only patchy prescription records for his claimed 22 months of use. The jury also found that Merck failed to warn both men of the drug's cardiovascular risks, and that it committed consumer fraud by misrepresenting those risks to doctors. The verdicts are the first on the long-term use of Vioxx (see *Nature* 440, 277; 2006).

Germany approves cash boost for high-tech science

The German cabinet has given the go-ahead for a €6-billion (US\$7.3-billion) programme to strengthen academic and industrial research and development. The biological sciences, nanotechnology, information technology and space technology are among the priorities of the government's high-tech strategy, announced on 5 April.

The extra money will bring Germany a step closer towards spending 3% of its gross domestic product on science, says Annette Schavan, Germany's science minister.

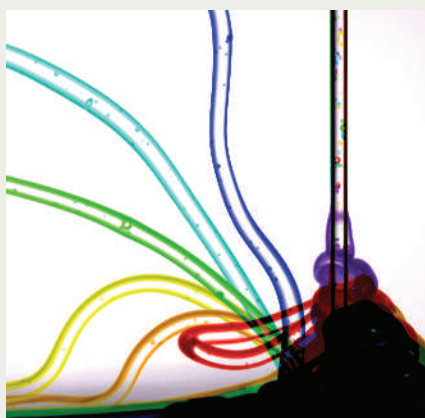
The windfall begins this year with an injection of €700 million into science — including €280 million for universities and academic research organizations — and will peak in 2009 at €2.2 billion.

Bouncing fluids make a splash

Dutch researchers believe they have cracked the physics behind a mysterious bouncing behaviour of liquids, called the Kaye effect. It occurs when viscous liquids are poured on to a surface; the down-going stream suddenly throws up a jet that merges with the incoming stream.

The Kaye effect was thought to be a strange property of complex liquids. But now, Michel Versluis of the University of Twente in the Netherlands and his colleagues say the effect may actually happen with everyday fluids, such as tomato ketchup, shampoo and liquid soap.

The team studied the Kaye effect using a high-speed video camera and found that it occurs in liquids whose viscosity drops as they flow. This property, termed shear-thinning, creates a slippery layer between the stream of liquid and the pool forming below, which allows the stream occasionally to jet up and away from the pool (see picture, which combines



several snapshots over a few hundred milliseconds). "It's a very simple experiment; you can do it in the bathroom," says Versluis (for videos see doi:10.1038/news060403-10).



D. WILSON/MISSOURI BOTANICAL GARDEN

GARDENS IN FULL BLOOM

In a world of declining biodiversity, botanical gardens are coming into their own — both as storehouses of rare plants and skills, and increasingly as centres of molecular research. **Emma Marris** reports.

At this moment there are a couple of ladies, one in her sixties, one in her eighties, walking through a greenhouse, their silver heads surrounded by hanging orchids in oranges, scarlets and lavender-tinged whites. These ladies, orchid-fancying mother and daughter, are bound to be there, because they were there in every garden visited during the writing of this feature. They are the embodiment of the botanical garden as popularly imagined: cosy, slightly old-fashioned, detail-oriented, perhaps a touch eccentric. But these greying dryads do not tell the whole story.

Tucked behind the palmhouses of the world's larger gardens are buildings filled with industrial wheeled shelves for mounted plant specimens, scientists sitting round-shouldered over microscopes, and growth chambers housing rows and rows of seedlings.

Botanical gardens have always been repositories of knowledge as well as cuttings. For centuries they were the heart of botanical scholarship. These days, most plant scientists work in academic settings, and increasingly study plants at a molecular level not obviously suited to the setting of greenhouses and flowerbeds. The challenge for

botanical gardens is to maintain a place in the scientific world while remaining true to their hybrid heritage, a heritage that encompasses aesthetics, exploration and education as much as academic study.

The two traditional scientific specialities of botanical gardens are plant taxonomy — the discovering, naming and sorting of species — and whole-organism biology — the study not of the ecosystem of the wild tulip, or the cellular functions that the tulip shares with a bunch of other plants, but of the tulip itself. Now, several of the world's larger gardens are broadening their focus and undertaking the sort of molecular investigation more typically found in research universities. Some botanists worry that this move stretches resources that would be better focused on the gardens' traditional strengths.

But researchers within the gardens argue that a botanical garden, filled with so much knowledge of plant diversity and bristling with plant life, is the ideal place to do such work. As head of molecular systematics at the Royal Botanic Gardens in Kew, southwest London (see 'Kew Gardens'), Mark Chase pores over DNA and RNA to probe plant relationships. But he'll still pop out on a spring

morning to look at the garden's many irises. He calls it "reciprocal illumination". Looking at the plant, then at the molecules, then back at the plant, "you really see things you hadn't noticed previously," he says.

Michael Donoghue is a university plant systematist who, although firmly rooted at Yale University in New Haven, Connecticut, understands the lure of working in gardens. He argues that molecular systematics are "part of the gardens' mission to understand plants in all their glory". The New York Botanical Garden's grand foray into genomics — the \$23-million Pfizer Research Center is opening there next month — may "raise some eyebrows", Donoghue says, "but why not endeavour to do all you can?" He adds that his counterparts at gardens are able to tackle much meatier projects for longer periods: "We go from grant to grant; at the gardens, they have a longer view and really devote a lot of time and energy to one mission."

Keeping up traditions

But Richard Olmstead, a molecular systematist at the University of Washington in Seattle, echoes the worry in the botany world that such moves could diminish the gardens' focus on taxonomy and whole-organism biology. "Insofar as any of their research is diverted towards more modern approaches away from those more traditional approaches, those traditional things are not going to get done, because no one is picking them up," he says.

No one is picking them up because there is, at the moment, no money in it. "Many people follow the money down to the molecular end," says Peter Crane, director of Kew, referring to the way that botany is often practised in academia these days. "It's a bit of a shame. We need a more integrated plant science." Edward Schneider, president of the Botanical Society of America, and a scientist at the Santa Barbara Botanic Garden agrees: "We need to preserve the understanding that a plant is more than just a bag of genes."

Membership in the American Society of Plant Taxonomists has been steady over the past ten years. Nevertheless, there does seem to be a want of expertise on plant diversity in the average university department. "The number of people does not seem adequate to deal with the challenges that



J. PEDEN/NEW YORK BOTANICAL GARDEN

Good spot: a new research centre means the New York Botanical Garden will provide a site for genomics as well as for picnics.

face the field," says Peter Raven, president of the Missouri Botanical Garden in Saint Louis (see 'Missouri Botanical Garden', overleaf). When universities want to compare species to deepen their understanding of a molecular finding, he argues, they increasingly turn to botanical gardens for help.

Some think that universities have abandoned the whole-organism approach. Dennis Stevenson, vice president for botanical science at the New York Botanical Garden, has a guess as to why: "Their faculty can't go running off into the jungle." In contrast, many botanical gardens' researchers start their projects with a vigorous mosquito-slapping tour through the remote and fecund places of Earth in search of specimens. "Most of the world's plant-diversity specialists are in the north, and most of the plant diversity is in the south," explains Crane. The most tangible fruits of these trips are large numbers of dried plants mounted on cards and bits of DNA in tubes filled with silica gel.

As well as being an insurance policy in the event of rare plants going extinct, the DNA samples feed into molecular

Kew Gardens

Founded: Kew Gardens became a botanic garden in the modern sense in 1759.

Research budget: £8 million (US\$14 million).

Famous for: The Palm House (pictured right), designed by Decimus Burton and Richard Turner in the 1840s; the ten-storey pagoda, built in 1762, which is open to visitors this spring and summer; and its winter orchid show.

Peter Crane (pictured right), director of Kew Gardens, is very neat and tidy — just a small, decorative ink stain on his cuff. He is walking the 129 hectares these days with a bit of preemptory nostalgia, as he leaves for the natural history Field Museum in

Chicago in the autumn. "I don't think once you've been here, you ever really get rid of the infection," he says. "We've planted literally millions of bulbs in my time here."

In the lake outside the Palm House, moorhens and huge white-headed geese float about. In the spring, carpets of bluebells, daffodils and crocuses are around every turn. In the corner of the garden are the research labs, a recent paper from which reported that two species of palm tree had diverged despite sharing



KEW

Peter Crane, head of Kew, cherishes botanists' traditional skills in whole-organism biology.

the same habitat (V. Savolainen *et al. Nature* doi:10.1038/nature04566; 2006).

E.M.



KEW

work. The mounted plants end up in herbaria — libraries of many millions of plants that are consulted by scientists. Although many university departments still have herbaria, they are increasingly becoming concentrated in botanical gardens. “Botanical gardens are becoming depositories of herbaria, because universities are moving in a more molecular direction,” says Schneider.

Jim Solomon runs the Missouri herbarium. He says that when he feels anxious or stressed he likes to sit down and sort a box of uncatalogued specimens. The plants last a surprisingly long time, attached to paper backing with linen strips or Elmer’s glue, and, although often rather brown, can retain a few lively characteristics. A maple specimen — *Acer macrophyllum* — collected in 1892 by Emma Shumway in Seattle, Washington, looked as if it had been caught falling from the tree just seconds ago. “As habitats are modified, as plants go extinct, these collections will become increasingly important as a record of what these things looked like,” says Solomon.

Plant protectors

Although preserving the vanished is important, preserving the all-but-vanished is even more so. Botanical gardens support conservation in two ways. They study plant diversity and establish what is rare where. And they grow rare plants and keep their seeds. *Cosmos atrosanguineus*, the chocolate cosmos plant, can now be smelt and seen only in gardens such as Kew, because it is extinct in its native Mexico. Unfortunately, all the chocolate cosmos plants alive today seem to be cuttings of a single plant — so they are ‘self-incompatible’ and won’t reproduce. Preserving genetic diversity is a large part of the gardens’ agenda. “There is an increased focus on not just having one individual with a label on it, but creating *ex situ* populations of threatened plants,” says Pete Hollingsworth, director of genetics and conservation at the Royal Botanic Garden Edinburgh.

The London-based Botanic Gardens Conservation International (BCGI) keeps some 600 member gardens in touch with each other on conservation matters. “I would say that the vast majority of our members have some sort of research programme,” says Suzanne Sharrock, director of



“As plants go extinct, these collections will become increasingly important as a record of what these things looked like.” — Jim Solomon

public awareness and understanding at BCGI. “In some cases it would be a very small programme, but the fact that they are botanical gardens rather than just public parks indicates that they are trying to be something more than a nice place to have a picnic.” Botanical gardens in Australia, South Africa, Brazil, and, increasingly, China, are praised by botanists from Europe and North America for their garden science and conservation activities. Sharrock thinks that if they all spoke with one voice, that voice might be pretty loud. “There are botanical gardens in every country in the world, and in every major city,” says Sharrock. “They have the potential to be a very powerful force.”

At the Second World Botanic Gardens Congress in 2004, the delegates adopted 20 goals for 2010 in support of the Convention on Biological Diversity’s Global Strategy for Plant Conservation. Among these goals is doubling the number of “trained botanical-garden staff working in conservation, research and education”. Another is compiling “a working list of known plant species”.

Various bits and pieces of such a possible database already exist. The International Plant Name Index strives for comprehensiveness in cataloguing the seed plants of the New World and Australia, but it only lists names. Raven thinks that his garden’s database of vascular plants (pretty much everything except mosses and liverworts), TROPICOS, might be “an important part of a final strategy”, pointing users to published work on the species, and in some cases to maps of ranges.

Plans for the ultimate database inevitably lead to talk of DNA barcoding. If species-specific differences in defined DNA sequences were matched with a species name in some kind of database, an untrained person could use a sequencer or a DNA-chip to read the barcode in a botanical sample, send it to the database, and get back a name and all other necessary taxonomic data.

Apart from its undoubted geeky appeal, such a technology would in principle save a lot of time and drudgery. Carrying out identifications for colleagues at home and round the world is time consuming and uncompensated. The use of

Missouri Botanical Garden

Founded: 1859. Still operated under the terms of founder Henry Shaw’s will.

Research budget: US\$9.4 million.

Famous for: The Climatron geodesic dome and free concerts on Wednesdays in the summer.

Peter Raven is the president of the Missouri Botanical Garden, a 32-hectare spread in Saint Louis. With an egg-shaped head and an engaging smile, he has lived in a corner of the garden — like a vicar in his vicarage — since 1971. However his status as, among other things, the lead author of the bestselling plant-sciences text for undergraduates gives him an influence far beyond the bounds of his parish.

“Botanical gardens have always been scholarly enterprises,” Raven says, but the Missouri garden is an

aesthetic one, too, managing neatly to be attractive even in the harsh winters of the US Midwest. The Japanese garden, for example, looks best in the snow, its distinctive lines highlighted with white.

On the paths, school kids in khakis are punching each other. Under the osage orange trees are nets to keep the large green fruit from “bonking people on the head”, according to Lisa Brandon, the public-relations manager. In this garden, like others, everything is emblazoned with the name of some donor or other. Some science research in gardens is funded by a similarly direct route, which means no peer-reviewed proposals and very little red tape. Just a thank-you note.

E.M.



Visitors flock to the Missouri Botanical Garden to see the Climatron dome.

J. JENNINGS / MISSOURI BOTANICAL GARDEN

barcoding would free up people to do their own research.

But Raven is cautious about such a scheme. He thinks that many lifetimes could be consumed setting the system up — time that would be spent making lists rather than learning anything about the organisms. “There are many millions of nematode species,” he says, choosing an animal example. “If I had a slide of every nematode in the world, what would I do with it?” Similarly, what would one do with barcodes for the 13,000 or so moss species?

Gardener's tips

There is, after all, so much else to look at, especially when one has the intellectual freedom that a large botanical garden can provide. Their widespread sources of funding and, in some ways, more forgiving range of stakeholders to please, mean that gardens can consider long projects and quirky studies that universities would be hard pushed to take on. In Missouri for example, 42% of the funds allocated to research comes from tickets, memberships and sales, and 58% comes from government grants, private donors and foundations. You can be “more creative,” says Ken Cameron, an orchid specialist and head of molecular systematics at the New York Botanical Garden. “At a university, you don't really have a choice. You have to toe the line and study one of these model organisms.” He prefers to study “weirdo little orchids that nobody cares about”.

At the gardens, researchers also enjoy a great deal of public support. Amateur gardeners want to know how to keep their plants alive and blooming, and look to the professionals for help, as well as for inspiration. They are also often curious about the microscopic details and ecosystem-level stories behind plants. This makes botanical gardens ideal forums for fostering the public understanding of science. The Royal Botanic Garden Edinburgh is raising funds for a building designed specifically to mediate interactions between scientists and the public. This will open “with a bit of luck” in 2009, says Stephen Blackmore, regius keeper of the garden. “A lot of people want to hear about the research,” says Blackmore, adding that scientists in botanical gardens seem to be regarded as “a good source of reliable, factual information” on everything from climate change to genetically modified organisms.

The exchange of information is not all one way. Amateur gardeners can have valuable information to impart. “For many gardeners, the information is in their heads. The tricks they use to propagate plants — how to grow a pineapple in Cornwall — don't always get written down,” says Crane. The value of this knowledge has only just been realized. Crane says he sometimes looks at the churchyard of Saint Anne's, opposite Kew's main offices, and thinks of the lost skills buried with generations of gardeners there.

Raven feels that amateur gardeners could do more to keep rare and endangered plants alive. “There is probably more scope for scientists to get involved with gardening. I could see a lot of room for home gardeners maintaining genetic diversity in a world that's becoming more homogenous.” And why not have them do research too? Some botanical gardens, including the one in Edinburgh, are already trying to track the subtleties of climate change by comparing various plant milestones year on year. Bringing home gardens into such networks would greatly increase the geographical reach of Edinburgh's researchers. Raven even imagines gardeners being issued with genetically identical indicator plants to make the data set really tidy.



Leafy retreat: Peter Raven, head of the Missouri Botanical Garden, thinks gardeners could be key players in efforts to protect biodiversity.

“Botanical gardens are really the only places that have the skills to adapt the landscape to changing conditions.”

— Stephen Blackmore



Gardens may thus have a functional role in the struggle to understand environmental change. But as important or more so, say Crane and Blackmore, is their inspirational role: the model they provide of how to relate to the flora of the Earth. Metaphorical thinking about the plant world has swung like a pendulum over the past decades. A hundred years ago, humans saw plants as resources to be deployed in ways that best served man — whether in amber fields of grain or in formal strolling gardens. With the rise of environmentalism, the view that humans should let nature run its course and the wild run wild has gained strength.

But the wild, these days, is rather piebald. Roads and wires and concrete interrupt it almost everywhere. Some of the major constituents in various ecosystems are all but gone. So to keep what's there, it may be necessary to actively care for what is left, rather than to leave it be. “Most ecosystems are not what they were, and they have to be managed,” says Crane. And care is the hallmark of the gardener.

Blackmore says that only “long-term thoughtful intervention” will protect plant diversity. And in the future, with climate change increasingly apparent and familiar cycles out of whack, only a competent, calm cadre of scientific gardeners may be able to tell the world how to keep the plants we rely on going, he says. “Botanical gardens are really the only places at the moment that have the skills to adapt the landscape to those changing conditions. Maintaining biodiversity in the face of climate change is going to be a very active process.”

Emma Marris is a Washington correspondent for *Nature*.

P. HOWARD/MISSOURI BOTANICAL GARDEN

Kew

BREAKING NEW GROUND

In 1906, a great earthquake destroyed San Francisco, and galvanized US seismologists. **Naomi Lubick** looks back at the event that changed the country's geological scene.

The Great Earthquake and subsequent fire that destroyed San Francisco in 1906 began at 5:12 a.m. on 18 April. More than 3,000 people are thought to have died following the magnitude-7.9 tremor. The metropolis of San Francisco, built on gold-rush fortunes, was almost utterly destroyed in three days of fire, and officials spent years playing down the possibility of another 'big one'. Yet the earthquake also jump-started seismology in the United States, inspiring it to catch up with countries such as Britain, Japan and Germany.

The US scientific community had already encountered several major earthquakes. Three tremors of magnitude 8 or more racked the New Madrid region in the US Midwest in 1811 and 1812. And the city of Charleston, South Carolina, was seriously damaged during an 1886 earthquake.

But the 1906 earthquake happened in the right time and place to act as a catalyst for science. Chance brought together several ingredients: the right people, the right technology, key ideas in need of testing — and a huge earthquake delivering the data. "It took that large an event to make seismology a national priority," says Jack Boatwright, a seismologist at the US Geological Survey (USGS) in Menlo Park, California.

Around the world, the discipline of seismology began to coalesce in the late nineteenth century, as non-specialists interested in earthquakes began to work together. In India, Richard Dixon Oldham identified primary (P) and secondary (S) waves, the main components of seismic waves. In Germany, engineers developed precision techniques to make seismometers more accurate. British geologists teaching in Japanese universities became interested in the tremblings beneath their feet. And in 1891, when the great Nobi earthquake hit, Fusakichi Omori of Tokyo's Imperial University was primed to lead his colleagues in documenting the event.

The Japanese view the magnitude-8 Nobi earthquake in much the way that Americans see the 1906 San Francisco event, says Thomas Jordan, director of the Southern California Earthquake Center in Los Angeles. Thousands of people died, and supposedly 'modern' structures were shaken apart. In tracking the subsequent tremors, Omori made critical observations that led to his eponymous theory on how aftershocks decay with time. And he brought his ideas with him to San

Francisco, where he did field work after the 1906 earthquake.

Omori's expertise was rare in the city. "In my estimation, there weren't any seismologists in the United States before 1906," says Boatwright, "there

were geologists." That soon changed. Immediately after the tremor, Andrew Lawson, a professor of geology at the University of California, Berkeley, assembled a troupe of scientists, who fanned out across the affected area. The group became a kind of seismic SWAT team, collecting evidence north towards Oregon and south almost to Mexico.

Ground forces

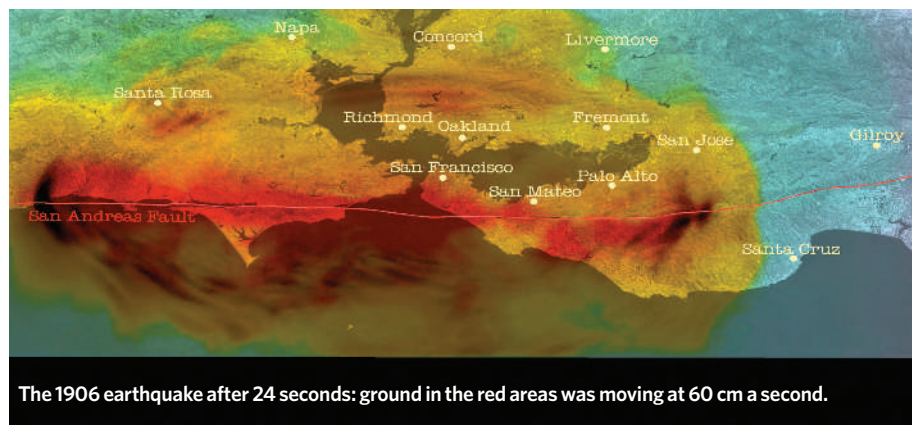
For the first time, scientists began to recognize the extent of the San Andreas fault. This 1,300-kilometre gash marks the boundary between

"There weren't any seismologists in the States before 1906."
— Jack Boatwright



two sections of Earth's crust: the Pacific plate and the North American plate. California geologists knew of the fault's existence, but until 1906 they had little idea of the power it could unleash. They did, however, realize that such faults cause earthquakes — a theory that came in part from Omori's documentation of the Nobi disaster.

Hours after the earthquake rattled San Francisco, geologists saddled up, got into automobiles — recently introduced to the region



G. K. GILBERT/USGS

CREATED BY USGS FOR 100TH ANNIV. EARTHQUAKE CONF.



Crack shot: Grove Karl Gilbert's photo of a rupture is part of a 1908 report on the 'big one'.

— or simply walked to sites along the San Andreas fault. "It was a really remarkable group of people," says Boatwright. Participants ranged from the respected geologist François Matthes, who avoided talking to anyone, to the sociable Grove Karl Gilbert, a USGS legend who spent a lot of time chatting to people who lived along the fault and recording their personal experiences. Gilbert quickly picked up evidence of the fault's movement, noticing fences, houses and roads that had been offset by up to 6 metres in some places.

Stirring tale

In 1908, Gilbert, Lawson and their colleagues produced a large tome that included detailed descriptions of fault scarps, sag ponds and other evidence of the earthquake; it was accompanied by an atlas of maps and seismic profiles. The Lawson report was reprinted in 1969¹, and seismologists and geophysicists still use it today for its physical descriptions, maps and timetables of the 1906 event.

The scientists who created it didn't know

what they were looking for, so they documented the earthquake observations very simply, in terms of factors such as building damage, fault movement and behaviour of the soil, says Dave Wald of the USGS in Golden, Colorado. In part, these extremely detailed descriptions are what makes the Lawson report so useful today.

Putting the whole picture together was not easy. At the time there was no overarching theory of plate tectonics — how parts of Earth's crust grind against each other. So geologists couldn't quite figure out how the San Andreas fault worked. "People didn't have a mechanism for large-scale deformation," says Carol Prentice of the USGS in Menlo Park.

Gilbert and others had already noted that faults could move horizontally, not just vertically. And geologists working on the Lawson report could clearly see that the fault had ruptured for at least 430 kilometres along its length, without significant uplift. Eventually, the San Francisco event helped geologists to recognize that the San Andreas is a strike-slip fault, in which two plates slide past each other rather than moving under or over one another.

All in store

In 1910, a second volume of the Lawson report was published, written by Harry Fielding Reid. It contained the 1906 event's greatest and most lasting contribution to seismology: the theory of elastic rebound. A professor of physics and geology at Johns Hopkins University in Baltimore, Maryland, Reid proposed that faults store up stress until they can no longer hold it, at which point they snap like a rubber band stretched too far. His meticulous synthesis came from field observations and information about the timing of shaking along the San Andreas fault.

Before the San Francisco earthquake, Reid had worked almost exclusively on glaciers, particularly on how they advance and retreat over time. Apparently, he saw a similar pattern in earthquakes. Elastic rebound implied that faults would gradually store stress over time, rupture in an earthquake and then begin the process all over again. And that, for the first time, suggested that earthquakes recur regularly on the same fault, in potentially predictable cycles.

The USGS estimates that an earthquake similar to the 1906 event is not likely in the near future, although earthquake risk remains high in the region — particularly on the

nearby Hayward fault. San Francisco residents, city planners and emergency officials are currently bracing themselves for an earthquake of magnitude 6.7 or greater, which has a 62% probability of occurring in the San Francisco Bay area².

This preparation is a lesson that California has taken decades to learn. In a scenario that may seem distressingly familiar to disaster planners today, Lawson struggled after the earthquake to get funding for follow-up studies and to spread knowledge about seismic risks. Members of his team helped to found the Seismological Society of America in August 1906, modelling themselves on Japan's Seismological Society. But, says Duncan Agnew, a historian of seismology at the University of California, San Diego, "for a long time 1906 was one of the few well-documented earthquakes."

Landscape view

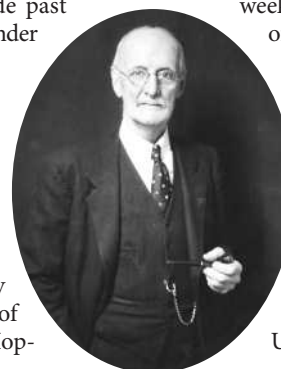
Even now, scientists are coming up with more analyses of the long-past earthquake. At next week's centennial anniversary meeting of the seismological society, Boatwright and his USGS colleagues will unveil a 'shake map' for the 1906 event. To produce a map of the intensity of shaking, they combed through the Lawson report to recreate the rupture as it would have been documented by seismometers placed along the fault.

Geologist Tina Niemi of the University of Missouri in Kansas City will present the meeting with a reconstruction of earthquakes along the section of the fault that jumped the farthest, in

Marin County. She and her co-authors will report that large earthquakes recur on that part of the fault every 50 to 600 years, so some activity might be due soon.

On the same day, the few survivors from 1906 will gather, as is customary, at 5:12 a.m. at Lotta's Fountain in downtown San Francisco, to lay a wreath of remembrance and tell stories. Thousands are expected to attend, as San Francisco embraces the legacy of its most devastating day.

Naomi Lubick is a science writer in Washington DC.



Harry Fielding Reid posited that tremors come in cycles.

1. *The California Earthquake of April 18, 1906: Report of the State Earthquake Investigation Commission* (Carnegie Institution, Washington DC, reprinted 1969).
2. Working Group on California Earthquake Probabilities *Earthquake Probabilities in the San Francisco Bay Region: 2002–2031* (US Geol. Surv., 2003). <http://pubs.usgs.gov/of/2003/of03-214>

BUSINESS

University challenge

Cambridge is revamping its approach to technology transfer — but will it work? Colin Macilwain reports.

The University of Cambridge is surrounded by a high-technology boom. The area is one of the fastest-growing and wealthiest in Europe, with an estimated 1,500 high-tech companies employing some 45,000 people. And in the popular imagination, this success is widely attributed to the transfer of ideas from the university into businesses.

But until recently, the university took a surprisingly relaxed approach to technology transfer. Academics were allowed to develop their own patents pretty much as they saw fit, and the university's income from royalties was fairly low by international standards: some £2.7 million (US\$4.7 million) last year. The region's high-tech success actually came about without any real concerted effort by the university itself.

That is about to change. Last December, the university revised its rules to give it first option on the patent rights for any discovery made in its labs. And Cambridge Enterprise, the unit established to run technology transfer and help young firms at the university, will be spun off later this year as a free-standing company. It will be run by Teri Willey, a US technology-transfer specialist with a strong track record in leading research universities in the US midwest. Her brief is to boost royalty income while providing a level of service for academics that will match the best in the world.

For Cambridge academics, the changes have come not a moment too soon. "The university has been pretty poor" at supporting technology transfer, says one academic with extensive industrial interests, who declined to be identified. "There's been too little support, and it's been poorly focused."

Originally, the technology-transfer office was part of the university's research-services division. But a review in 2004 led to Cambridge Enterprise being split off from the division, and the unit is now set to gain much more autonomy.

Fully owned by the university, the new company will negotiate licensing agreements for patents held by the university — although academics will be allowed to go elsewhere with any patents that it doesn't license out. It will also advise academics on setting up their own companies, provide some



M. MNISZKO/UNIV. CAMBRIDGE

Cambridge is seeking to foster innovation beyond the ancient university's walls.

of them with seed capital, and help them source larger amounts of venture capital.

The company's mission is to serve the academics, stresses Ian Leslie, pro-vice-chancellor for research. "It is not a vehicle to make vast amounts of money for the university," he says. Nonetheless, he says that top US schools aim to generate 4–6% of their research income from licensing fees on their patents. "At Cambridge, we're running at 1 or 2%," he says. "I think over ten years we can catch up."

The grand design for Cambridge Enterprise nevertheless got off to a slow start. After it split from the research-services division, the unit

struggled to secure a permanent director — largely, according to university officials, because neither the reform of intellectual-property rules nor plans for the new company were complete.

Willey will arrive in August to take the helm, and she is enthusiastic about the unit's prospects. "There's a very solid foundation for doing this kind of work," she says. She describes the climate for licensing and venture capital

in Cambridge as slightly tougher than the hotspots on the US coasts. "But it's actually very similar to the areas where I've worked," she says. "There are a few big venture-capital funds and lots of developing ones." Nevertheless, Willey notes that there is a gap between what hap-

pens in Britain and at the top US environments in Boston and northern California.

In 2004, the Cambridge cluster attracted about 8% of all European venture-capital investments. "It's definitely one of the best clusters in Europe," says David Gill, director of Cambridge-based ET Capital. "But the downside is that it is actually quite small, especially compared with Silicon Valley." Some £500 million in fresh capital has been committed to Cambridge companies in the past decade, Gill says — considerably less than the amount raised last year alone in the area around Stanford University in California.

Economists say that the university's technology-transfer office is only one small component in the network of relationships that help high-tech businesses get off the ground. And some academics remain sceptical about its role. "I don't actually believe that the technology-transfer efforts have that much impact on how these things happen," says Henning Sirringhaus, a Cambridge physicist and co-founder of electronics firm Plastic Logic. Sirringhaus and others argue that entrepreneurial expertise inside academic departments counts for more.

Leslie nonetheless sees Cambridge Enterprise as a vital cog in the larger, university-industrial complex. "The University of Cambridge is now part of an innovative cluster that relies on an entrepreneurial culture," he says. "Cambridge Enterprise is a part of that, and it's a part that we'd like to see work very well." ■



Teri Willey: a strong track record in the US midwest.

C. M. HANSON

Industry: policing the 'dark side' of ecology

SIR — I commend your News Feature “Caught between shores” (*Nature* **440**, 144–145; 2006), highlighting the rift between academic and corporate ecology. I fully support your call for a higher standard of ecological science to regulate business activities and defend the environment. However, considering that environmental issues are gaining global and political centre stage, and with a growing awareness of the need to preserve natural heritage, I feel you stopped short of speaking out on the truly crucial issues at hand.

There has been considerable movement at governmental level during the past few decades to implement a host of environment-protection legislations. These are designed to force businesses to consider the ecological implications of their actions within a legal framework, usually through an environmental impact assessment (EIA).

But there has been insufficient back-up or policing of these policies. The major blunder is that the responsibility for organizing the ecological studies required for an EIA is left to the very companies who are supposedly being regulated. It may well be that the science performed within such companies is sound and impartial. But without an official system to regulate the EIA process, is it really surprising to see hostile attitudes among ecologists in academia towards colleagues perceived as going to the ‘dark side’?

Perhaps a United Nations-sanctioned professional body should be created to govern scientists involved in EIA preparation. The Ecological Society of America runs a professional certification scheme that would be a useful model for such an action. Or perhaps EIAs should, by law, be outsourced to ‘regulated’ ecological consultancies.

If big businesses have a genuine ethical policy, they will support such actions. If they don’t, the wheat will be sorted from the chaff.

David Allsop

School of Biological Sciences AO8,
University of Sydney, NSW 2006, Australia

Industry: speak up to stop its pressure on academia

SIR — Your News Feature “Caught between shores” (*Nature* **440**, 144–145; 2006) suggests that scientists in industry sacrifice independence for influence. So do scientists in government. It’s a classic insider/outsider dilemma. As a former chief environmental scientist for Australian Mineral Development Laboratories, I’ve tried both sides. But companies and governments make no claim to put truth before profit or politics. Universities do.

A more severe and insidious threat to scientific integrity is when governments link research funding to industry involvement, or when universities hire people with a background in commercial cut-and-thrust rather than academic ethical practice.

What to do? As a scientist, you can keep your independence if you are cautious with unsolicited commercial contacts and careful with contracts, and if you have no dependents, so you can ignore threats — whether financial, legal or worse. You can keep your reputation if you publish in good journals, maintain competitive research funding as well as industry contracts, work (for free if need be) for community watchdogs on the industry concerned, and take part in public debate and expert advisory councils. Just don’t expect much spare time, or influence.

At a societal level, academic independence would be served by better separation between sectors: industry for entrepreneurship, government to regulate, universities for knowledge. This would work best with a fourth entity to develop and apply university research for industry and government and to insulate academics from commercial and political pressure. Some government research organizations used to do just that, but now their roles, too, have become blurred.

Currently, peer support is our best hope. Scientific and professional societies have a critical role. If your colleagues are being pressured improperly, help them. Your turn won’t be far away.

Ralf Buckley

Environmental and Applied Sciences,
Griffith University, Parklands Drive, Gold Coast,
Queensland 9726, Australia

Risks of a high-protein diet outweigh the benefits

SIR — Alastair Robertson of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) says that the CSIRO’s high-protein Total Wellbeing diet is “based on peer-reviewed science within robust experimental frameworks” (“Diet’s healthy blend of science and practicality” *Nature* **439**, 912; 2006). These small studies reported no significant difference in weight loss between a high-protein meat-based diet and a control diet with lower protein content. The exception was a small sub-group of women with high triglyceride levels, who lost more weight over 12 weeks with a high-protein diet.

Longer-term trials of high-protein diets are more controversial, but some studies by the CSIRO and others show that such results do not last, and that weight loss and sustainability are not superior to diets that focus on a reduction in fat and overall energy intake (see G. D. Brinkworth *et al. Int. J. Obes.* **28**, 661–670; 2004). Robertson’s claim that the

Total Wellbeing diet can “contribute to reducing obesity in Australia” is hype, not science. The diet is not a more viable option than current dietary recommendations.

Recent cohort and laboratory studies (T. Norat *et al. J. Natl Cancer Inst.* **97**, 906–916; 2005, and M. H. Lewin *et al. Cancer Res.* **66**, 1859–1865; 2006) also highlight the potential increased risk of colorectal cancer with a high intake of red and processed meat — both prominent in the CSIRO diet. Add the high financial and ecological costs of diets high in meat, and they are not justified in the absence of any superior weight-loss benefit.

Rosemary Stanton*, Tim Crowe†

*School of Medical Sciences, University of New South Wales, Randwick, NSW 2031, Australia

†School of Exercise and Nutrition Sciences, Deakin University, 221 Burwood Highway, Burwood, Victoria 3125, Australia

Local people may be the best allies in conservation

SIR — We, like many, have been excited by the discovery of new animal and plant species in West Papua’s Foja Mountains. Although we are not against granting the area official protection status, as discussed in your News story “Calls to conserve biodiversity hotspots” (*Nature* **439**, 774; 2006), we warn against imposing such schemes on local people.

Protection status in itself is no panacea: elsewhere in Indonesia deforestation inside protected areas often outpaces that outside (see L. M. Curran *et al. Science* **303**, 1000–1003; 2004). But there is considerable scope for arrangements that respect local claims and interests while also benefiting conservation goals. Working with Indonesian partners and Conservation International in the Mamberamo-Foja region, we have found that locals are valuable allies for conservation. Indeed, they have been solely responsible for protecting the Foja until now, and it was the local people who made the recent Foja expedition possible. They are sensitive about these ancestral lands, and have driven off outsiders seeking minerals and other resources in the past. But, once a firm basis for trust has been established, they provide enormous input, creating maps of special sites and resources with their traditional knowledge.

Local communities must not be viewed as a problem, but as central to the solution.

Douglas Sheil*†, Manuel Boissière*‡

*Center for International Forestry Research, PO Box 6596 JKPWB, Jakarta 10065, Indonesia,

†Forest Ecology and Forest Management, Wageningen University, PO Box 47, 700 AA Wageningen, The Netherlands

‡Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Campus de Baillarguet, 34398 Montpellier Cedex 5, France

BOOKS & ARTS

Real concerns, false gods

Invoking a wrathful biosphere won't help us deal with the problems of climate change.

The Revenge of Gaia: Why the Earth is Fighting Back — And How We Can Still Save Humanity

by James Lovelock

Allen Lane: 2006. 192 pp. £16.99

Tyler Volk

James Lovelock, the renowned scientist who pioneered the measurement of atmospheric trace gases, here offers his take on the future of energy. In brief, a vigorous turn towards nuclear power will be necessary to prevent the catastrophic climatic changes caused by an increase in atmospheric carbon dioxide, a by-product of burning fossil fuel.

Several decades ago, Lovelock developed an idea called the Gaia hypothesis, named after the ancient Greek goddess of the Earth. He postulated that “life on Earth actively keeps the surface conditions always favourable for whatever is the contemporary ensemble of organisms”. Evolutionary biologists such as Ford Doolittle and Richard Dawkins countered, by showing that neither life as a whole nor specific living things could be selected by evolution to create favourable global surface conditions, and Lovelock has conceded this point. He soon developed what he now calls Gaia theory.

What is Gaia and what is the theory? Gaia is the system consisting of all life and the environments that life affects — soils, atmosphere, oceans and surface rocks. The theory states that Gaia is a closely coupled, evolving system that has the goal of keeping the climate favourable for life as it currently stands. This idea of a goal parts company with how virtually all Earth and environmental scientists see the world, and is important here because Lovelock's new book blends Gaia theory and metaphor with a viewpoint on energy and the greenhouse effect.

The concerns are truly massive. Our planet is in crisis. Rising atmospheric carbon dioxide is affecting the biosphere's radiative balance, ecosystems and ocean acidity. According to Lovelock, we have inadvertently declared war on Gaia, and she “now threatens us with the ultimate punishment of extinction”. I submit that such anthropomorphic language does more harm than good to the author's arguments. But more on that later.

In *The Revenge of Gaia*, Lovelock evaluates the candidates for energy sources that do not emit carbon dioxide. Spreading inherently



IMAGE STATE/ALAMY

Making peace with Gaia? Nuclear power stations can provide energy without destroying the biosphere.

huge networks of solar, wind or biomass farms only further desecrates the vital and increasingly scarred landscape of the biosphere and should be limited to small doses. Carbon sequestration deserves trials but the scales required are stupefying. The only recourse is the large-scale deployment of nuclear fission, at least as a stopgap until nuclear fusion and an adequate portfolio of renewables comes along.

Lovelock compares the tiny volumes of nuclear waste with the mountains of annually emitted carbon dioxide. He answers those he sees as fear-mongers about radiation and provides daunting numbers for renewables, such as the quarter-of-a-million wind turbines that would be required to power the United Kingdom. He is a systems thinker. His informed, often iconoclastic ideas are, as usual, worth pondering. There are gems here. For instance, we need to work out the technicalities to let us prepare to deploy macro-engineering fixes, such as sun-blocking reflectors in space, just in case the biosphere goes into a runaway mode of uncontrollable warming.

But beware the use of Gaia theory in all this. Lovelock says that he takes the goddess as a metaphor and a thinking aid, like a ship's captain referring to the vessel as ‘she’, and admits that Gaia theory is “provisional and likely to be displaced”. These caveats, however,

are overwhelmed by the unquestioned presence of Gaia.

Quite simply, the problem is that Gaia as portrayed here is false. Essentially the same criticism levelled by evolutionary biologists against the Gaia hypothesis holds for Gaia theory. There can be no goal for the biosphere, which, however extraordinary, is an understandable chemical mixture of air, water, soil and organisms. This system settles naturally into steady states, bounded by local negative feedback, and the current states are dear to our lives and worthy of protection. But Lovelock persists in using the language of goals and even intentions for the biosphere system.

For instance, he states that Gaia buries organic carbon “to keep oxygen at its proper level”. Wrong. Gaia doesn't do things for reasons. The organic carbon that slips past the bacterial layers in ocean sediments has global effects, but such effects are not why the carbon is buried. The incorporation of carbon by plants and the emission of cloud-enhancing gases by marine plankton did not evolve as climatic air-conditioning mechanisms, as in another of Lovelock's discussions; they evolved because they enhanced the internal metabolic needs of specific organisms.

We are also told many times that Gaia likes it cool. But the biosphere (my preferred name

for Gaia) doesn't like or dislike anything. So why does Lovelock persevere with this language and mode of thinking? Why doesn't he let his analysis of energy systems and feedbacks in climate and the carbon cycle stand on its own? Lovelock is aware, with regard to his personalization, of the criticisms of scientists. Yet he claims to be unrepentant.

Certainly the language of goals and intentions gives emotional juice to the arguments. But when Lovelock promotes a metaphor that mixes in false science, the metaphor itself is tainted and must be dismissed.

He makes a case that metaphor is needed. But to pronounce that Gaia is "angry" with us, or that she will "eliminate those that break her rules" weakens the otherwise rational (albeit

controversial) scrutiny of our complicated global jam. Too often, Gaia here seems less like science and more like one man's mythology elevated into the service of deeply felt environmental concerns. Lovelock likens the incomprehensibility of Gaia to that of God, valorizes those who declare allegiance to Gaia, and claims that Gaia theory is a seed from which an "instinctive environmentalism can grow; one that would instantly reveal planetary health or disease and help sustain a healthy world". What does "instantly reveal" mean?

Read this book for its thoughtful sections on global energy and climate, but steer clear of its web of Old Testament-like prophecy. ■
Tyler Volk is in the Department of Biology, New York University, New York 10003, USA.

enough to detect this influence — large mountains can affect readings strongly enough to yield mapping errors of more than a mile.

Danson's book revolves around an expedition to Mount Schiehallion in the Scottish Highlands, which he interprets as a key founding event in the science of geodesy. Organized by the Royal Society, an underpaid team set out with the finest available equipment, but tempers flared during long weeks of arduous labour. Sleeping in cramped tents and hampered by atrocious weather, astronomers and surveyors quarrelled about their duties, and personality clashes were further exacerbated by class differences — although the astronomer royal, Nevil Maskelyne, redeemed himself by sending a Stradivarius to the ghillie whose fiddle had been burnt during the farewell celebrations. Back in London, Maskelyne calculated Schiehallion's attraction by comparing the astronomical and land measurements of its width. Armed with this result, and thanks to the surveyors' astonishing feat of triangulating the whole mountain to find its volume, he estimated its density.

This is a handsome book, brimming with original images as well as modern diagrams, although someone at the publisher apparently believes that the secondary in "secondary sources" means less important. An alert editor might have prevented Danson from dating experiments to test Newton's gravity to 15 years before he published the *Principia*, and also toned down phrases such as "Mother Nature was not about to let her foolish children exploit her oblate body." But this modern surveyor's love for both his craft and his historical material emerges strongly. Danson not only explains technicalities clearly, but also gives readers some notion of the endurance and attention to detail involved in drawing maps before satellite systems removed the hard work — and the romance — from surveying.

Danson is a narrative rather than an analytic author, who intersperses scientific discussions with ample background information of varying relevance. His generosity in reproducing the minutiae of his research unfortunately obscures the plot-line leading to today's theories of Earth's structure, and this emphasis on presenting facts precludes a discussion of their significance, which would have given his work a wider appeal. Although the dust-jacket blurb highlights the novelty of Danson's narrative, his text skates around the interesting question of why Maskelyne's strenuous activity on Schiehallion has been forgotten, whereas Michell and Cavendish have acquired canonic status. To say they got a better answer is too simple: Danson's work implicitly illustrates how the theoretical work of scientists was — and still is — more esteemed than the practical contributions of engineers. ■

Patricia Fara is at Clare College, Cambridge. Her recent books include *Newton: The Making of Genius* (2002) and *Fatal Attraction: Magnetic Mysteries of the Enlightenment* (2005).

Climbing Mount Schiehallion

Weighing the World: The Quest to Measure the Earth

by Edwin Danson

Oxford University Press: 2006. 289 pp.
£17.99, \$29.95

Patricia Fara

The Enlightenment recluse Henry Cavendish is remembered for being eccentric, synthesizing water and determining the Earth's density. Notoriously taciturn, Cavendish apparently made one joke in his entire life: commiserating with the ill-health of his colleague John Michell, he hoped that "it may at least permit the easier and less laborious employment of weighing the world". Their efforts with a torsion balance eventually proved successful, but these heroes of mensuration make surprisingly brief appearances in *Weighing the World*. Instead of describing their laboratory experiments, Edwin Danson focuses on a far less

famous line of investigation — measuring the attraction of mountains.

Danson, a civil engineer, has coined a misleading title for his book, which turns out to be mainly a history of surveying in Britain and India during the 'long eighteenth century'. He starts his account in seventeenth-century France, when perplexing results from accurate instruments forced cartographers to challenge the long-standing assumption that God had made the Earth a perfect sphere. They concluded that the Earth must be slightly squashed — but, as Voltaire put it, battles raged about whether it resembled a melon or a lemon. Several expeditions later, it seemed that the Earth was more like a pear. Surveyors faced an additional challenge to accuracy. Isaac Newton himself pointed out that, according to his gravitational theory, mountains must be centres of attraction. But Newton was wrong to doubt that instruments would ever be sensitive



Weighing up the evidence: Mount Schiehallion was the focus of an early attempt to measure the world.



Thirsty work

Nearly a billion people live in the drylands of the developing world. These often isolated mountain and desert communities spend much of their time trapping, channelling and conserving what little water there is. As a result, they are expert in one of today's urgent tasks — learning to cope with finite resources.

Dry: Life Without Water (Harvard University Press, £22.95) tells 16 stories of dryland life, from fog catching above Chile's Atacama Desert to the forest nurseries helping to regenerate Burkina Faso's near-barren terrain (pictured). A distillation of work by the Academy of Sciences for the Developing World and the Third World Network of Scientific Organizations, these snapshots of sustainability are models of how science and traditional knowledge can profit from each other.

The riveting images, from Massai managing their wetlands to vicuña shearers at work in the high Andes, underline a message that is vital at a time when drought, desertification and impending water wars are high on the international agenda.

B.K.

Weird, sexy and mind-blowing

The Quantum Zoo: A Tourist's Guide to the Never-Ending Universe

Marcus Chown

Joseph Henry Press: 2006. 216 pp. \$24.95

Jim Al-Khalili

This entertaining account of modern physics carries an interesting subtitle. Although I am not sure what is meant by "Never-Ending" — infinite in physical extent, or in intellectual scope? — I agree with the phrase "A Tourist's Guide", as that is exactly what this book is. It is an entertaining little gem that leads the reader through many of the wonders of twentieth-century physics with a light and sometimes quirky touch that I thoroughly enjoyed. It is so full of little insights and neat analogies that I found myself folding over the top corners of countless pages containing quotable passages. I have seen many of the descriptions and analogies elsewhere, but that is to be expected. What is remarkable is the number of new ways Marcus Chown has found to explain difficult and often abstract concepts.

The Quantum Zoo is really two books in one. The first part deals with the quantum world, whereas the second focuses on Einstein's special and general theories of relativity. The latter subject is easier to discuss at a non-technical level, and Chown seems more relaxed once he has guided us through the counter-intuitive and wacky quantum world. He stomps through quantum superposition, interference,

uncertainty, tunnelling and indistinguishability with breathless enthusiasm. On consecutive pages he refers to the "weirdest manifestation of [quantum] entanglement", the "sexiest potential use of entanglement", and the "most mind-blowing consequence of entanglement". Such superlatives carry the reader through the subatomic landscape, not always appreciating what one sees — as is inevitable with quantum mechanics — but being left in no doubt that this really is a mind-blowing subject.

I preferred the second half of the book, as I felt that the author enjoys the subject matter more. Here we move to the einsteinian scale of the very fast, the very large, and curved space-time. Chown returns to the quantum only in the last two-and-a-half pages, when he touches on theories of quantum gravity.

I should give one example of the author's eye for a quirky analogy. There are many ways to describe the effects of gravitational energy, but in his playful and understated style, Chown describes the force with which a falling roof slate hits you on the head. He suggests that you measure gravitational energy via the heat generated from the impact. He then asks how much energy would be released were the gravitational pull of the Earth trillions of times stronger — at which point I found myself worrying more about the size of the bump on my head. Great stuff.

I could not fault the physics in either half of

the book, and would take issue with only a couple of points. In his description of special-relativistic time dilation Chown describes how, if you were to travel very close to the speed of light, your time would slow down so much that the Universe's whole future would flash by in an instant. Of course, this would be true only if you were speeding up or slowing down. If you were travelling at constant velocity, you would see the opposite happen: everything outside your rocket would run in slow motion. After all, motion is, well, relative.

Chown also jumps the gun when claiming that NASA's Gravity Probe B experiment has confirmed the general-relativistic effect of space-time frame dragging by the Earth's rotation when, as far as I am aware, the data are still being analysed. I am not in any doubt that the effect will be observed and reported, probably in the coming year, such is our confidence in the predictions of general relativity. But there is always a hint of doubt — otherwise why do the experiment?

Such quibbles aside, I hope this book does well. It is a great introduction for anyone who does not care to gain a deep understanding of the subject, or sign up to an author's pet theories, but just wants to know what all the fuss is about. This is what good popular science writing is all about.

Jim Al-Khalili is in the Department of Physics, University of Surrey, Guildford, Surrey GU2 7XH, UK.

Eggs and exegesis

Putting the 'history' back into natural history.

Martin Kemp

Eggs and Easter are now indissolubly linked. But like many of the Christian rituals, the origin lies in other cultures. From China to Egypt, the egg is the symbol of rebirth in spring, and even of the Universe itself. No natural object is more loaded with meaning. The wonder of its shape and the miracle of the life within have an endless fascination.

The largest egg of all, the ostrich's, has drawn intense secular and religious attention. Prized in cabinets of curiosities, the eggs were often given extravagant mounts in silver. In churches they hung in places of honour. The most famous is suspended from the shell niche behind the Virgin in Piero della Francesca's Brera Altarpiece (pictured here), painted in the mid-1460s for Duke Federigo Montefeltro in Urbino.

In the thirteenth century, Bishop Guilelmus Durandus gave two reasons for exhibiting ostrich eggs in church. Like other exotic curiosities, such as unicorn's (really narwhal's) horns, they testified to the wonder of creation. More specifically, they carried an intricate Christian message, derived from the *Physiologus* or *Bestiary*, probably compiled in the fourth century from classical, Egyptian and Indian sources. It told of the ostrich's strange incubation habits.

A "careless" creature, says the *Physiologus*, the "sparrow camel" (*Strutho camelus*) abandons its eggs in the sand, but returns to hatch them by staring at them. The heat of its sight incubates the chicks, in accordance with the theory that seeing rays emanate from the eye. The Christian interpretation given in the *Physiologus* and by subsequent writers is that worshippers should keep their eyes on Christ at all times.

These legends may seem to have little to do with what we regard as the science of animals. However, Konrad Gesner, the great Swiss natural historian often seen as the father of zoology, on the basis of the superb illustrations in his *Historia Animalium* (1551), diligently records stories from the *Physiologus* and subsequent bestiaries. It seems that we may have incorrectly characterized the origins of modern science.

The learned text supplied for each illustrated beast begins by comparing the animal's names in various languages, a necessary preliminary in the era before standard nomenclature. Then Gesner deals with the animal's appearance and form, its food, generation and life in its habitat, including its enemies. He next outlines its utility for man, before an extensive section on its wider history in various cultures, its symbolic, moral and geographical



THE ART ARCHIVE/GALLERIA BRERA MILAN/DAGLI ORTI

The appearance of an ostrich egg suspended behind the Virgin in Piero della Francesca's Brera Altarpiece indicates its symbolic importance to the Church.

connotations, and its role in literature.

The aim was nothing less than a comprehensive review of each animal in its widest context from a humanist perspective. His account emphasizes the 'history' in natural history — a name that now applies largely to museums — in the face of the rise of the more 'scientific' terms zoology and botany, or, more recently, animal and plant sciences. A significant aspect of the story of the modern sciences of nature consists of combating what came to be seen as the literary, artistic and antiquarian stigmas associated with natural history. Not the least

of the avoiding actions was to abandon the deluxe picture-book that had dominated the study of animals and plants from the sixteenth century.

But the legends of the bestiary remain deeply embedded, albeit largely unnoticed, in our own culture. We talk, for instance, of crying crocodile tears, as crocodiles were reputed to weep insincerely after devouring a man. And, of course, if we do not want to face reality, we bury our heads in the sand like an ostrich.

Martin Kemp is professor of the history of art at the University of Oxford, Oxford OX1 1PT, UK.

NEWS & VIEWS

ASTROPHYSICS

A whirling dervish

Richard Gray

Vega is a fundamental reference star for astronomers. But it seems that our perceptions of it have been misconceived — rather than spinning slowly, the star is a rapid rotator seen pole-on.

On page 896 of this issue, Peterson *et al.*¹ present measurements that convincingly show that Vega — the second brightest star in the northern night sky — is not turning unhurriedly as had been assumed, but rotating at a speed near to that at which it would fly apart. Confirming this behaviour required cutting-edge technology, as Vega's rotational axis is pointing almost directly at Earth. The discovery would seem to demand a rethink of some established ideas — among them, our notions of the distribution of energy in the electromagnetic spectrum of astronomical objects, and perhaps even of the time required for planets to form around stars.

For astronomers, Vega is arguably the most important star in the sky. In the standard Morgan–Keenan system, which classifies stars into types O, B, A, F, G, K and M in order of decreasing temperature, Vega is the archetypal A-type star. These stars, with temperatures around 10,000 kelvin and with a bluish-white tinge, are among the most common stars visible to the naked eye. Vega is also the fundamental 'spectrophotometric' star, serving as the standard calibrating source of light for telescopes working in regions of the electromagnetic spectrum from the far ultraviolet through the visible to the infrared. In addition, astronomers have historically used Vega — which is an apparently simple, slowly rotating star — to test and develop their theories of stellar atmospheres and interiors. Finally, the star is the archetype of a class of stars with disks of dust and debris surrounding them that are possible sites of planetary formation.

For some decades, however, there has been an uneasy feeling that all is not right with Vega. In 1964, the star failed to conform to R. M. Petrie's model² that allowed the intrinsic luminosities of most hot stars to be calculated from the strength of the hydrogen lines in their spectra. Vega was found to be too luminous by a factor of nearly two, leading Petrie to suggest that it was actually two stars, too close to one another to be resolved by a telescope. Because no Doppler shifts — changes in the observed frequency of radiation from moving objects — were observed in Vega's spectrum, Petrie

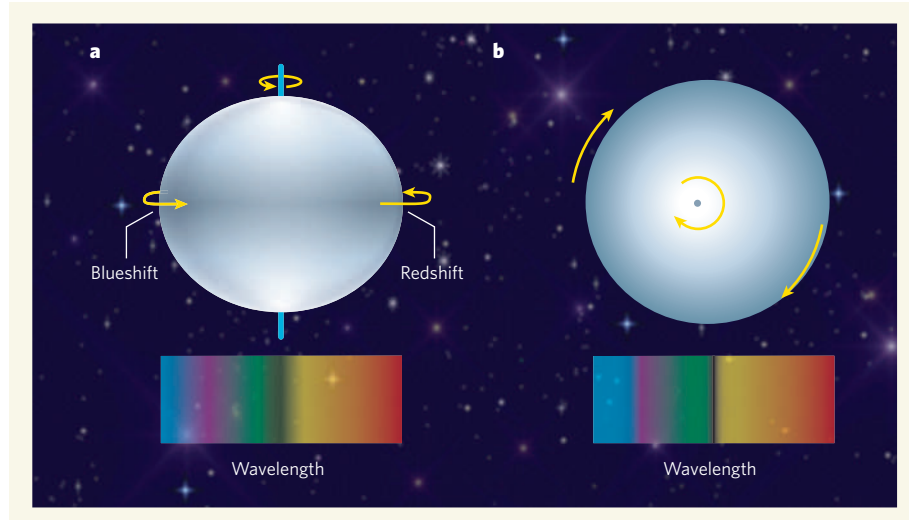


Figure 1 | Doppler broadening. Elements in a star will absorb photons at characteristic wavelengths, leading to absorption lines (drops in light intensity) in the spectrum. **a**, The Doppler shift that occurs in frequency when the velocity of an object relative to an observer changes, as with the tone of a passing ambulance siren, also crops up in stellar observations. Radiation from a rapidly rotating star seen face-on will be shifted to shorter blue wavelengths on the side rotating towards us, and to longer red wavelengths on the side rotating away. The result is a rotational broadening of the expected absorption lines. (The broadening effect is magnified here; in reality, it amounts to only a few tenths of a nanometre, even for the most rapidly rotating stars.) **b**, In contrast, little or no variation in velocity, and so no Doppler broadening, will be discerned in a star seen along its axis of rotation. Slowly rotating stars look similar, as the difference in perceived velocity on the sides rotating towards and away from us is also small. Vega, assumed from its lack of Doppler broadening to be a slowly rotating star, has only now been outed^{1,7} as a rapid rotator seen almost pole-on.

speculated that our line of sight must be exactly perpendicular to the orbital plane of the double, so that, as they orbit one another, neither star is ever moving towards or away from Earth.

Three years later, however, intensity interferometry revealed that Vega is a single star³. In the ensuing decade, it was demonstrated theoretically⁴ that a rapidly rotating star viewed in its equatorial plane would seem fainter and redder than a slow rotator, whereas a rapid rotator seen pole-on would be unusually bright. This phenomenon occurs because a rapidly rotating star is flattened by the greater centrifugal force at its equator, which has the effect of making the surface temperature there significantly cooler than at its polar regions.

In rapidly rotating stars seen in their

equatorial plane, Doppler shifts from the high rotational velocities both towards and away from us — on opposite sides of the star as we look at it — combine to smear out absorption frequencies and produce broadened spectral absorption lines (Fig. 1a). Slowly rotating stars and rapid rotators seen pole-on, in contrast, have narrow spectral lines, as the variation of rotational velocity along our line of sight is small (Fig. 1b). In all respects except its anomalously high luminosity, therefore, a rapid rotator seen pole-on mimics a slow rotator.

Following on from this insight, I proposed in the mid-1980s that Vega's anomalous luminosity arises because it is rotating near its break-up velocity, but is seen nearly pole-on⁵. In 1991, highly detailed spectra of Vega⁶ revealed that, although the absorption lines

are narrow (like those of a slow rotator), they have unusual shapes. These line shapes could be explained only if Vega were rotating rapidly with its axis of rotation inclined to our line of sight by about 5°.

This was a marvellous, but indirect, confirmation of my original proposal. Since then, powerful techniques to study stars directly have been developed. These include optical interferometry, in which telescopes separated by as much as hundreds of metres are combined optically to form one enormous 'compound' telescope. The advantage of this technique is that it leads to exquisitely high resolution, revealing stars as disks, rather than mere points of light, and allowing details on their surfaces to be discerned.

Peterson *et al.*¹ used such an instrument — the Navy Prototype Optical Interferometer in Flagstaff, Arizona — to study Vega, and show that the distribution of light over its surface is slightly asymmetrical. According to their model, this light distribution is consistent with Vega rotating at 93% of its break-up velocity, with its polar axis inclined at 4.5° to our line of sight. Simultaneously, another group has confirmed the pole-on, rapid-rotator model for Vega using the CHARA interferometer on Mount Wilson near Los Angeles, California⁷.

These results^{1,7} have profound implications for astronomy. Vega, our prime astronomical calibrator, clearly has a surface temperature that varies from about 10,000 K at its poles to only 8,000 K at its equator. As some parts of our calibration of Vega's spectrum assume a slow rotator with a uniform temperature over its surface, this will affect our use of Vega as a 'standard lamp' and thus our derivations of the spectral energy distributions of astronomical objects from asteroids to galaxies.

Vega is also manifestly not a simple, slowly rotating star as had been thought. Alternative models for the make-up of Vega are needed, together with updated estimates for the abundances of metals in its interior. Because determining the age of a star depends on its measured metallicity, that might in turn require the revision upwards of estimates for the age of the star, perhaps from the current 350 million years to around 570 million years.

Our understanding of Vega's debris disk — the closest to Earth known, and therefore the one that has been most thoroughly studied — will have to be overhauled, too. As the disk lies in Vega's equatorial plane, it is experiencing a much cooler radiation field than was formerly assumed. That fact must be borne in mind when accounting for the infrared radiation that the disk re-emits, and might affect our estimates of its density, as well as the size of particles that dominate it. If Vega's age is indeed greater than assumed, models of how long debris disks survive also need re-examination. Planets could be forming in Vega's disk right now, so that would have a knock-on effect on ideas about the time required for the formation of planets. ■

Richard Gray is in the Department of Physics and Astronomy, Appalachian State University, Boone, North Carolina 28608, USA.
e-mail: grayro@appstate.edu

1. Peterson, D. M. *et al.* *Nature* **440**, 896–899 (2006).
2. Petrie, R. M. *Publ. Dominion Astrophys. Obs. Victoria* **12**, 317–338 (1964).

3. Hanbury Brown, R., Davis, J., Allen, L. R. & Rome, J. M. *Mon. Not. R. Astron. Soc.* **137**, 393 (1967).
4. Collins, G. W. III & Sonneborn, G. H. *Astrophys. J. Suppl. Ser.* **34**, 41–94 (1977).
5. Gray, R. O. *J. R. Astron. Soc. Can.* **82**, 336–348 (1988).
6. Gulliver, A. F., Adelman, S. J., Cowley, C. R. & Fletcher, J. M. *Astrophys. J.* **380**, 223–229 (1991).
7. Aufdenberg, J. P. *Astrophys. J.* (in the press); preprint available at www.arxiv.org/abs/astro-ph/0603327 (2006).

AGEING

Chromatin unbound

Jan Vijg and Yousin Suh

Sir2 proteins slow ageing in yeast by locking chromatin — the DNA and proteins in chromosomes — into a stable, silent state. Inactivating a Sir2 family protein in mice causes premature ageing and genome instability.

Genomes are exquisitely adapted to provide the information, at the right time and place, for cells to function. Loss of genome integrity has long been implicated in ageing: cancer, for example, occurs more frequently with age and can be caused by genome alterations. Writing in *Cell*, Mostoslavsky *et al.*¹ now show that, in mice, inactivation of SIRT6 — a protein related to Sir2, which increases lifespan in yeast and worms when overexpressed — causes genome instability and premature ageing.

Ever since the discovery in the late 1940s that low, daily doses of radiation accelerated the signs of ageing in rodents², loss of genome integrity has been proposed as a universal cause of ageing³. Under normal conditions, the most likely culprits for causing genome deterioration are the reactive oxygen species continually produced during metabolism. It comes as no surprise that the enzymatic systems that detect and repair DNA damage are widely considered candidate longevity mechanisms. Heritable defects in such genome maintenance processes are generally associated with premature ageing syndromes. Werner's syndrome, for example, is caused by lack of the WRN protein that is crucial for some forms of DNA repair and for maintaining the structural integrity of chromosome ends. In both humans and mice, WRN mutations and other defects in pathways that protect genome integrity cause premature ageing⁴.

Intiguing players in genome maintenance are the silent information regulator (SIR) proteins, which repress gene expression and genome instability in yeast by stabilizing chromatin. The founding member of this family, Sir2, promotes longevity in yeast by protecting regions in the genome that consist of repeated DNA sequences, which are inherently unstable. Moreover, extra copies of Sir2 increase the lifespan of fruitflies and nematodes⁵. Interestingly, in some species Sir2 has been connected to the 'pro-longevity' effect of both caloric restriction (that is, food deprivation without malnutrition) and the reduced insulin/insulin-like

growth factor-1 (IGF-1) signalling that is a response to nutritional stress but can also occur as a consequence of mutations in genes acting in the insulin/IGF-1 cascade⁵.

The actions of Sir2 are not straightforward and there has been no direct evidence to link Sir2 with ageing in mammals. Its closest mammalian relative, SIRT1, regulates the tumour-suppressor protein p53 and FOXO3 (a gene regulatory factor that is negatively regulated by insulin/IGF-1 signalling) to suppress programmed cell death (apoptosis) and promote cell survival⁶. Inactivating SIRT1 in mice is mostly lethal, with the few live births dying within weeks from abnormalities that do not resemble ageing⁷.

Mostoslavsky *et al.*¹ show that inactivation in mice of another Sir2 relative, SIRT6, quickly causes abnormalities that resemble some aspects of ageing, such as thinning of the skin due to loss of subcutaneous fat, and signs of osteoporosis. Cells from these mice showed impaired proliferation and increased genomic instability, possibly because of a defect in base excision repair (BER). This would be the first demonstration that accelerated ageing is associated with a defect in this DNA-repair pathway. BER is especially crucial for removing DNA damage caused by reactive oxygen species. It is unclear how SIRT6 promotes BER, but because Mostoslavsky *et al.* find that the protein is associated with chromatin, it may regulate accessibility of BER enzymes to the damaged site.

Premature ageing symptoms in mice with defects in DNA repair have been linked to cellular responses to the increased DNA damage, such as apoptosis and cellular senescence (the irreversible cessation of cell division)⁴. Mostoslavsky *et al.* found a dramatic increase in apoptosis of lymphocytes — cells active in the immune response — in SIRT6-deficient mice. This was the most likely cause of the severe lymphocyte depletion seen in the animals. Unexpectedly, however, bone marrow transplantation experiments indicated that

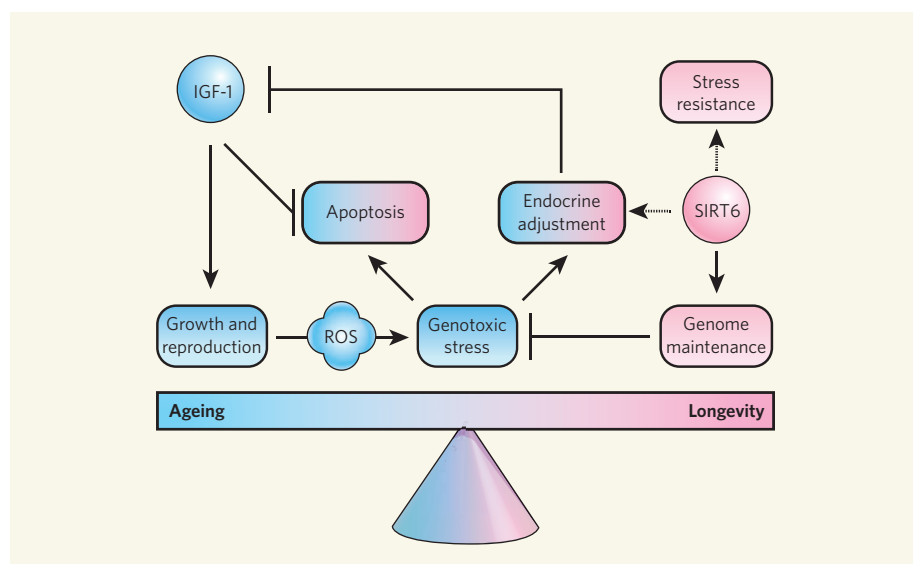


Figure 1 | Balancing act. The possible role of SIRT6 and genome maintenance in balancing ageing (blue) and longevity (pink). Although a mechanism is as yet unclear, SIRT6 may act in its pro-longevity role by promoting DNA repair, increasing stress resistance and maintaining metabolic homeostasis. The genotoxic stress resulting from its absence may cause a compensatory metabolic shift towards reduced insulin-like growth factor (IGF-1) signalling, thereby lowering the production of reactive oxygen species (ROS). The resulting increase in apoptosis could be a cause of the premature ageing symptoms observed by Mostoslavsky *et al.*¹ in mice lacking SIRT6.

the lymphocyte depletion was not cell-intrinsic but rather a response to a systemic defect. Analysis of the blood serum of SIRT6-deficient mice revealed extremely low levels of IGF-1, as compared with normal, control animals. IGF-1 strongly inhibits apoptosis in lymphocytes, and the age-related reduction in lymphocyte production by the thymus has been ascribed to a decline in IGF-1 levels with age⁸. This finding presents us with a paradox, because reduction of IGF-1 signalling is linked with a longer, not a shorter lifespan. How can these results be reconciled?

Life extension conferred by dampening IGF-1 signalling is likely to represent a metabolic switch in the use of resources, away from growth and reproduction with the inevitable side effect of DNA damage, towards increased maintenance and repair⁹. Although defects in insulin signalling in mammals cause diabetes, reduced IGF-1 signalling extends life in mice⁹. The attenuation of IGF-1 signalling in the SIRT6-deficient mice may indicate that this switch has been activated as a means to limit the onslaught of spontaneous DNA damage and to upregulate the apoptosis of severely damaged cells. Indeed, modulating IGF-1 signalling, possibly through members of the Sir2 family, may be a general mechanism for coping with environmental stress, including nutrient limitation and genotoxic stress. Interestingly, Xpd^{TTD} mutant mice, which harbour a defect in another form of DNA repair known as nucleotide excision repair, display characteristics of accelerated ageing and caloric restriction, suggesting that metabolism adjustment might be a general mechanism to deal with genotoxic stress (Fig. 1)¹⁰.

So does SIRT6 connect BER to the IGF-1

signalling pathway? As yet, this is far from clear. For example, neither the nature of the SIRT6 enzymatic activity, nor its potential substrates, is known. Indeed, Mostoslavsky and colleagues' results only indirectly implicate SIRT6 in BER, so although it is reasonable to assume that SIRT6 promotes the access of repair enzymes through chromatin remodeling, there is no direct evidence for this. Also,

it is not apparent if and how SIRT6 is involved in regulating the IGF-1 response. Finally, not all degenerative symptoms associated with reduced lifespan necessarily involve the causes that underlie natural ageing. Indeed, the degenerative symptoms observed in SIRT6-deficient mice are far from comprehensive and may well result from developmental defects.

Nonetheless, Mostoslavsky *et al.*¹ provide a potential link in the chain connecting genome instability, metabolic effects and ageing. Ultimately, this may lead to ways of exploiting the role of Sir2 family members in genome maintenance to silence senescence.

Jan Vijg is at the Buck Institute for Age Research, 8001 Redwood Boulevard, Novato, California 94945, USA.

e-mail: jvijg@buckinstitute.org.

Yousin Suh is in the Department of Molecular Medicine and Barshop Institute for Longevity and Aging Studies, University of Texas Health Science Center, 15355 Lambda Drive, San Antonio, Texas 78245, USA.

e-mail: suhy@uthscsa.edu.

1. Mostoslavsky, R. *et al.* *Cell* **124**, 315–329 (2006).
2. Henshaw, P. S., Riley, E. F. & Stapleton, G. E. *Radiology* **49**, 349–364 (1947).
3. Szilard, L. *Proc. Natl Acad. Sci. USA* **45**, 30–45 (1959).
4. Hasty, P., Campisi, J., Hoeijmakers, J., van Steeg, H. & Vijg, J. *Science* **299**, 1355–1359 (2003).
5. Vijg, J. & Suh, Y. *Annu. Rev. Med.* **56**, 193–212 (2005).
6. Motta, M. C. *et al.* *Cell* **116**, 551–563 (2004).
7. Cheng, H. L. *et al.* *Proc. Natl Acad. Sci. USA* **100**, 10794–10799 (2003).
8. Taub, D. D. & Longo, D. L. *Immunol. Rev.* **205**, 72–93 (2005).
9. Bartke, A. *Endocrinology* **146**, 3718–3723 (2005).
10. Wijnhoven, S. W. *et al.* *DNA Repair (Amst.)* **4**, 1314–1324 (2005).

CONSERVATION BIOLOGY

Roads and genetic connectivity

Jared L. Strasburg

The Ventura Freeway slices through wildlife habitat near Los Angeles. A case study shows how large highways such as this can seriously impede genetic exchange in large vertebrates.

Roads are bad news for wildlife. Vehicles kill animals attempting to cross them, sometimes in large numbers. Conservation biologists have recognized for many years that roads also split up populations, and they continue to explore the degree and consequences of this fragmentation. Techniques such as radio-telemetry and motion-triggered photography can be used to observe animals crossing highways, but not whether the animals mate on the other side. In a seven-year study published in *Molecular Ecology*¹, Riley *et al.* compare direct observations with genetic estimates of the disruptive effects of a busy highway in southern California. They show that the highway acts as a social barrier to gene flow in two large and

mobile carnivores, and that it does so to a much greater extent than would seem to be the case from observations of animal dispersal.

Many animals, especially vertebrates, alter their behaviour to avoid roads, shifting home ranges, feeding sites and nesting areas away from them². But most animal groups still attempt to cross these barriers with some frequency, often by making use of culverts and other potential crossing points^{3–5}. These crossing points may be specifically designed for animal dispersal, such as large vegetated overpasses or underpasses; or they may be unintentional dispersal corridors where highways cross rivers, canyons or other roads.

Direct estimates of dispersal rates across

roadways are often high enough to lead researchers to conclude that habitat connectivity is adequate to maintain genetic connections between populations. But although these estimates may seem encouraging, the studies concerned (refs 4 and 6, for example) generally don't include data on levels of gene flow. Combining information from radiotelemetry and genetics, Riley and colleagues¹ have measured both dispersal and successful gene exchange in bobcats (*Lynx rufus*) and coyotes (*Canis latrans*) across the Ventura Freeway. They conclude that, for these species, relatively high rates of dispersal across the freeway do not translate into genetic connectivity.

The Ventura Freeway is one of the largest and busiest highways in the United States: it is 10–12 lanes wide and carries 150,000 vehicles every day. It divides a crucial habitat corridor between the Santa Monica Mountains to the south and areas of more extensive undisturbed habitat to the north. Riley *et al.* captured 110 coyotes and 87 bobcats within dispersal distance of the highway, fitted them with radio-collars, tracked their movements for seven years, and genotyped most of them at seven microsatellite loci. These loci are highly variable DNA markers, distributed throughout the genome, that can be used to identify individual animals and infer genetic relationships among populations.

The authors found that 4.5% of coyotes and 11.5% of bobcats crossed the highway at some point during their study period. Assuming two-year generation times for each species, these numbers translate into respective migration rates of 1.3% and 3.3% per generation. But based on genetic variation at the microsatellite



Figure 1 | A coyote on the verge. Roads act as artificial territorial boundaries for wildlife.

loci, they found significant differentiation between populations on either side of the highway; estimates of effective migration rates based on these genetic data were 3–18 times lower than direct estimates of dispersal.

A likely explanation for this discrepancy is 'home-range pile-up' — the radiotelemetry data indicate that home ranges abut the highway but do not cross it, so that it acts as an artificial territorial boundary (Fig. 1). Animals with home ranges bordering the highway show much higher range overlap with other individuals than do those farther away, and ranges adjacent to the highway are also much smaller. Migrants, especially young migrants, entering into this situation will find it more difficult to establish and defend a territory, and find mates. This in turn probably results in fewer individuals remaining and lower reproductive

success of those that do. The sample sizes of migrants are small, but Riley *et al.* note that six of the ten bobcats that crossed the highway soon returned, and neither of the two females that did remain on the other side produced litters the following spring. Only one of five coyotes that crossed the highway stayed on the other side during the mating season, and most dispersers were under two years old and so were probably less able to establish a territory.

Road ecology has been described as a 'sleeping giant' among environmental biologists⁷: there are indeed very few data concerning genetic fragmentation in animal populations separated by roadways, and what data are available are not encouraging^{8,9}. Not all highways are as busy as the Ventura Freeway, and not all animals will behave like the bobcats and coyotes studied by Riley and colleagues¹. But if the discrepancy between dispersal and gene flow that Riley *et al.* have found is widespread, as seems likely, the effectiveness of highway dispersal corridors needs to be tested to see how well they meet their goal of ensuring genetic connections among populations. There is evidently also a great deal to learn about how widespread the phenomenon of home-range pile-up is, the degree to which territorial and non-territorial animals may be differently affected by barriers such as highways and railways, and the extent to which corridor strategies thereby need to be adapted.

Jared L. Strasburg is in the Department of Biology, Indiana University, 1001 East 3rd Street, JH 142, Bloomington, Indiana 47405, USA. e-mail: jstrasbu@indiana.edu

1. Riley, S. P. D. *et al.* *Mol. Ecol.* doi:10.1111/j.1365-294X.2006.02907.x (2006).
2. Trombulak, S. C. & Frissell, C. A. *Conserv. Biol.* **14**, 18–30 (2000).
3. Rodriguez, A., Crema, G. & Delibes, M. *J. Appl. Ecol.* **33**, 1527–1540 (1996).
4. Gloyne, C. C. & Clevenger, A. P. *Wildl. Biol.* **7**, 117–124 (2001).
5. Foster, M. L. & Humphrey, S. R. *Wildl. Soc. Bull.* **23**, 95–100 (1995).
6. Beier, P. *Conserv. Biol.* **7**, 94–108 (1993).
7. Forman, R. T. T. & Alexander, L. E. *Annu. Rev. Ecol. Syst.* **29**, 207–231 (1998).
8. Epps, C. W. *et al.* *Ecol. Lett.* **8**, 1029–1038 (2005).
9. Keller, I., Excoffier, L. & Largiadere, C. R. *J. Evol. Biol.* **18**, 90–100 (2005).

CHEMISTRY

A nose for sarin

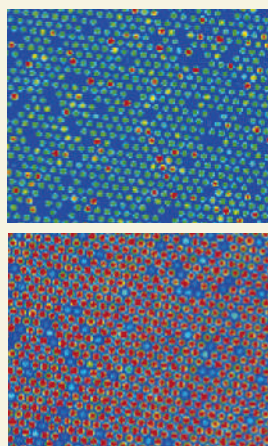
The gas attack on the Tokyo subway in 1995 demonstrated the need to rapidly detect and identify sarin and similar substances used in chemical warfare. David R. Walt and colleagues have devised a system to do just that (*J. Am. Chem. Soc.* doi:10.1021/ja057057b; 2006).

The authors attached a fluorescent indicator to polymeric, micrometre-sized beads. The indicator, fluoresceinamine, reacts with nerve agents such as sarin and soman that contain chemical groups known as phosphoryl halides. This causes the beads to fluoresce within seconds of a vapour burst (see lower image; red indicates fluorescence). As the reaction also generates acid,

which reduces the intensity of the fluorescence, the authors coated the beads with another polymer that included units of the organic compound pyridine. This basic layer neutralized the acid, maximizing the beads' shine.

The fluoresceinamine reaction is specific to a certain class of chemical, so the microbeads don't respond to other, chemically distinct warfare agents such as mustard gas. Crucially, everyday vapours such as ethanol, toluene and water also fail to stimulate fluorescence, allowing the unambiguous detection of sarin-like chemicals.

The beads can be easily integrated into current 'electronic nose' technology



used to detect and respond to many different vapours. Incorporated into the sensor array of such a nose, the sarin-detecting beads could react to small amounts of toxin in high concentrations of background vapours. Rapid-response nerve-agent detectors for public areas may soon be a reality.

Andrew Mitchinson

QUANTUM PHYSICS

Equilibrium on hold

Henk T. C. Stoof

Startlingly, two atomic clouds confined to one dimension can be made to pass through each other repeatedly without ever coming to rest. Such non-equilibrium phenomena are fundamental, but experimentally elusive.

As do the rest of us, nature likes to relax. Through microscopic collisions that result in macroscopic processes such as the diffusion of particles and energy, nature's first instinct is to seek a state of thermodynamic equilibrium, in which all mechanical, chemical and thermal variables reach constant values. In an ultracold atomic gas, for example, a state of equilibrium has generally been reached by the time each atom has collided with just three others.

The speed with which relaxation occurs makes observing non-equilibrium phenomena difficult. We are all, for instance, familiar with the equilibrium state of condensation on our windows on a cold day. However, actually seeing the non-equilibrium process — the formation and growth of the water droplets — is not so common. But help may be at hand: on page 900 of this issue, Kinoshita *et al.*¹ disclose that an atomic gas, when confined to an array of one-dimensional tubes by a web of laser beams, does not relax to equilibrium — even after each atom has collided several thousand times.

Physics in one dimension is often strange and counterintuitive. Two atoms brought very close to one another repel each other strongly, so, classically speaking, when an atomic gas is confined to a one-dimensional tube, every atom in the gas is restricted to oscillating back and forth between its two neighbours. This

situation can be brought about quantum mechanically too^{2,3}, in which case a so-called Tonks gas forms. In quantum mechanics, however, atoms can generally also pass through each other, and so do not have to be reflected completely. Both the fully and partially reflecting regimes are considered by Kinoshita *et al.*¹, with identical outcomes.

The results of their experiments are most easily understood by noting that, compared with the situation in three dimensions, a collision between two atoms in a one-dimensional tube is always very simple. In three dimensions, if two incoming atoms have a head-on collision, the outgoing atoms are located on a spherical shell that expands isotropically from the point of impact (Fig. 1a), as has been rather dramatically confirmed in experiment^{4,5}. Importantly, the collisions can change the velocities of the incoming atoms, allowing the velocity distribution of the gas to assume its equilibrium form — a process known as thermalization. If the collision takes place in a one-dimensional tube, however, the outgoing spherical shell collapses to two points, and the velocities of the outgoing atoms are exactly the same as those of the incoming atoms (Fig. 1b). Collisions cannot, therefore, change the velocity distribution of the gas and so bring about thermalization.

It is interesting to note here the subtle

difference between classical and quantum-mechanical collisions. Classically, we know for sure that the atoms have exchanged their velocities. Quantum mechanically, this is not true: a collision in which the atoms exchange their velocities interferes constructively with a collision in which the atoms do not exchange their velocities, and in fact pass through each other. The final outcome is, in both the classical and quantum cases, the same: after the collision, we have two atoms with the same velocities as before the collision.

Kinoshita and colleagues' experimental illustration¹ of this remarkable physics is beautiful in its simplicity. In a one-dimensional tube, atoms experience a 'harmonic' potential that grows in proportion to the square of their displacement along the tube from the central position. By using two counterpropagating laser beams to give one half of the atoms in the gas a kick in one direction, and the other half a kick in the other, two clouds of atoms can be created that oscillate exactly out of phase in the harmonic potential. Because the two atom clouds cannot exchange velocities through collisions and so thermalize, they will continue to pass through each other for ever.

This is a theorist's idealized view of the experiment. In practice, various complications require Kinoshita *et al.* to perform a careful analysis of their raw experimental data. One physically interesting complication is that, if two clouds pass through each other, an atom in one of the clouds can change its velocity — not through collisions, but because it feels a somewhat different average potential energy in the presence of the other cloud. This effect can, however, easily be avoided by giving the two clouds a sufficiently large initial velocity kick. Two further obstacles that had to be surmounted were the deviations of the potential from a truly harmonic form along the tube, and heating of the two atom clouds through collisions with the room-temperature atoms (which are always present, as the vacuum in which the experiment is performed cannot be perfect).

Kinoshita and colleagues' experiment¹ clears the way to exploring exciting new avenues in physics. One direction that springs to mind is tuning the web of laser beams that creates the array of one-dimensional tubes to change the nature of the atomic collisions from the one-dimensional to the three-dimensional case, and thus study how this affects the relaxation of the gas. Another promising path is to increase the temperature of the gas and so observe the growth dynamics of the state of matter known as a Bose–Einstein condensate in a controlled manner — approximately, the quantum analogue of observing the growth of a water droplet on a cold glass plate. Given the speed at which research into ultracold atoms is moving at present, there may be more suggestions. ■

Henk T. C. Stoof is at the Institute for Theoretical Physics, Utrecht University, Leuvenlaan 4, 3584 CE Utrecht, The Netherlands.
e-mail: h.t.c.stoof@phys.uu.nl

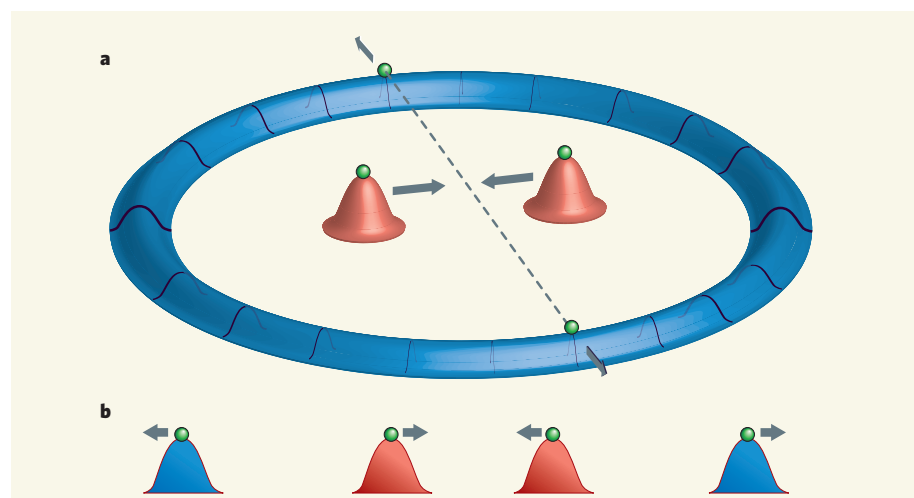


Figure 1 | Collision confined. **a**, Two incoming atoms (circles) and their associated quantum-mechanical probability densities (red) collide in three dimensions. Viewed from their centre of mass, the outgoing atoms will be equally probably anywhere on a spherical shell that expands at constant velocity from the point of impact. A slice (blue) through this shell is shown. **b**, Confined to one dimension, atoms can either rebound from or — quantum mechanically — pass through each other, with no change in the magnitude of the velocity (because of energy and momentum conservation) in either case. As Kinoshita *et al.*¹ show, with an appropriate potential, this behaviour repeats ad infinitum.

1. Kinoshita, T., Wenger, T. & Weiss, D. S. *Nature* **440**, 900–903 (2006).
2. Kinoshita, T., Wenger, T. & Weiss, D. S. *Science* **305**, 1125–1128 (2004).

3. Paredes, B. *et al.* *Nature* **429**, 277–281 (2004).
4. Chikkatur, A. P. *et al.* *Phys. Rev. Lett.* **85**, 483–486 (2000).
5. Buggle, Ch. *et al.* *Phys. Rev. Lett.* **93**, 173202 (2004).

BIOGEOCHEMISTRY

Methane and microbes

Rudolf K. Thauer and Seigo Shima

Microorganisms can carry out a wonderful range of chemical transformations. The anaerobic oxidation of methane seemed not to be among them. But it is — both with sulphate, and now it turns out, with nitrate.

Methane is not only a fossil fuel but also a key player in the carbon cycle. About 1% of the carbon dioxide annually fixed by photosynthesis is converted back to carbon dioxide by microorganisms via methane, which amounts to 1 billion tonnes of methane formed and consumed per year. Moreover, methane is a powerful greenhouse gas, and its concentration in the atmosphere has increased continuously over the past hundred years.

It is therefore news in itself that Raghoebarsing and colleagues have identified a hitherto unknown microbial sink for methane (page 918 of this issue)¹. Their discovery of the microbial consortium that mediates the process is also exciting — the microorganisms involved couple the oxidation of methane to denitrification (nitrate reduction to nitrogen) in the absence of oxygen (Fig. 1), which from a mechanistic viewpoint is a challenging reaction (alongside benzene, methane is by far the most unreactive hydrocarbon).

Raghoebarsing *et al.* isolated the methane-oxidizing microbial consortium from the anoxic sediments of a freshwater canal that was subject to agricultural run-off with a high nitrogen load (Fig. 2). The sediment was saturated with methane and contained nitrate at a concentration of 1 mM. The culture obtained grew on methane and nitrate (or nitrite) as the sole energy sources, with an estimated doubling time of more than one month. After 16 months, the culture mainly consisted of two different types of microorganism, a bacterium (80%) and an archaeon (10%).

Sequence analysis of the microorganisms' genomes revealed that the dominant bacterium belonged to a novel phylum without any documented cultured species, and that the archaeon was distantly related to archaea involved in the anaerobic oxidation of methane with sulphate. Similar sequences to those identified by Raghoebarsing *et al.*¹ have been retrieved from the denitrifying zone of freshwater sediments of Lake Biwa, Japan, and from contaminated groundwater in the United States¹, where methane disappears in the nitrate zone of the groundwater². So the anaerobic oxidation of methane coupled to denitrification is probably

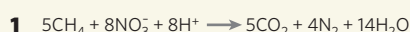


Figure 1 | The anaerobic oxidation of methane (CH₄). 1, The nitrate-dependent process identified by Raghoebarsing *et al.*¹ (reaction 2 in Fig. 2). 2, The sulphate-dependent process³. Reaction 1 involves a free-energy change of -765 kJ mol^{-1} . CH₄ has a specific rate of about 2 nmol per min per mg protein and an apparent K_M of $<1 \mu\text{M}$. The respective figures for reaction 2 are -21 kJ , 0.1 nmol per min per mg protein and $>10 \text{ mM}$.

of widespread ecological importance.

Why did it take until now to identify this process? The main reason is probably that it proceeds at much lower rates than the anaerobic oxidation of other organic compounds or of hydrogen sulphide with nitrate. So the process is evident only in anoxic environments with

low concentrations of oxidizable substrates other than methane, and with low levels of sulphate and high levels of nitrate. Such conditions prevail in anoxic freshwater habitats contaminated with agricultural run-off and in contaminated groundwater. The nitrate-dependent anaerobic oxidation of methane is predicted to occur mainly close to the oxic–anoxic interface, where in the oxic phase ammonia is oxidized to nitrate with oxygen by nitrifying (nitrate-forming) microorganisms. Nitrate then diffuses into the anoxic sediments saturated with methane, which is produced microbially from the degradation of cellulose and other plant material (Fig. 2). The interface is generally characterized by steep chemical gradients, which occur within millimetres and mask the process of the anaerobic oxidation of methane with nitrate.

The only other known anaerobic methane-oxidizing process is coupled to sulphate reduction³. This reaction is mediated by a consortium of methane-consuming archaea and sulphate-reducing proteobacteria⁴. The free-energy change associated with the reaction is much smaller than that for nitrate-dependent methane oxidation⁵. The two processes also differ markedly in their kinetics^{1,6} and in the apparent affinity for substrate (Fig. 1).

For the sulphate-dependent process, there is circumstantial evidence that a nickel enzyme, methyl-coenzyme M reductase, catalyses the first step of methane oxidation^{6,7}. The main evidence is that the methane-consuming archaea involved, which are related to methane-producing archaea, contain high concentrations of the enzyme — which, in the latter archaea, catalyses the actual methane-forming step. The catalytic efficiency of the

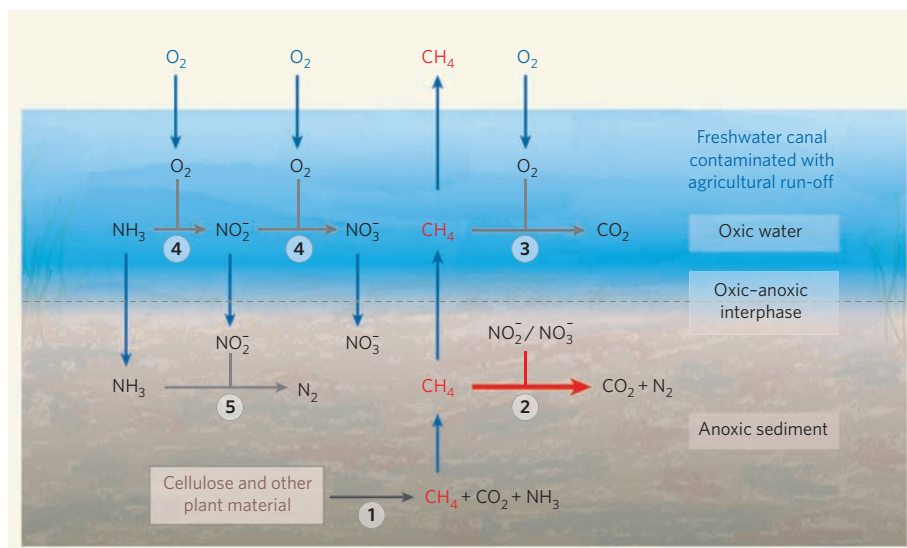


Figure 2 | Nitrate-dependent anaerobic oxidation of methane¹ and related processes. Methane oxidation is coupled to denitrification (NO₃⁻ and NO₂⁻ reduction to N₂) in fresh water that receives agricultural run-off and so contains high concentrations of ammonia (NH₃), nitrate (NO₃⁻) and nitrite (NO₂⁻). The sediment is anoxic and is saturated with methane. The blue arrows indicate transport by diffusion. The different processes (grey and red arrows) are mediated by different microorganisms. 1, Anaerobic bacteria, protozoa and methane-producing archaea. 2, Archaea related to methane-producing archaea and bacteria belonging to a new phylum (reaction 1 in Fig. 1)¹. 3, Methane-consuming bacteria. 4, Nitrifying bacteria. 5, Planctomycete bacteria.

enzyme in mediating the back-reaction might just be sufficient to account for the observed rates of sulphate-dependent anaerobic oxidation of methane in microbial mats⁸. But it is hardly consistent with the much higher specific rates and the much lower apparent K_M (Michaelis constant) for methane of the nitrate-dependent process.

The mechanism of the anaerobic oxidation of methane therefore remains unclear. But with the discovery made by Raghoebarsing *et al.*, a new door opens on the study of how methane is activated in the absence of oxygen.

Rudolf K. Thauer and Seigo Shima are at the

Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Strasse, D-35043 Marburg, Germany.

e-mail: thauer@mpi-marburg.mpg.de

1. Raghoebarsing, A. A. *et al.* *Nature* **440**, 918–921 (2006).
2. Grossman, E. L., Cifuentes, L. A. & Cozzarelli, I. M. *Environ. Sci. Technol.* **36**, 2436–2442 (2002).
3. Valentine, D. L. & Reeburgh, W. S. *Environ. Microbiol.* **2**, 477–484 (2000).
4. Boetius, A. *et al.* *Nature* **407**, 623–626 (2000).
5. Thauer, R. K., Jungermann, K. & Decker, K. *Bacteriol. Rev.* **41**, 100–180 (1977).
6. Krüger, M. *et al.* *Nature* **426**, 878–881 (2003).
7. Hallam, S. J. *et al.* *Appl. Environ. Microbiol.* **69**, 5483–5491 (2003).
8. Shima, S. & Thauer, R. K. *Curr. Opin. Microbiol.* **8**, 643–648 (2005).

used by the vesicles for identification, docking and fusion.

By labelling a vesicle-membrane protein called synaptotagmin with a fluorescent tag and using STED microscopy, Jahn and colleagues could follow the fusion of individual vesicles at the plasma membrane (Fig. 2 on page 937). Their study provides some of the most compelling evidence to date that at least some membrane constituents remain grouped together after vesicles fuse with the plasma membrane (rather than diffusing freely within the membrane like a drop of water on water), which is consistent with the kiss-and-run theory.

Although STED is not the first technique to break the diffraction limit, it is arguably the one best suited for biological imaging. For nearly two decades, near-field scanning optical microscopy has yielded images with a spatial resolution down to about 15 nm. Scanning is usually achieved by allowing light to escape through a tiny aperture in a metal-coated tip^{5,6}. Materials with negative refractive indices also hold promise for ultra-high-resolution optical imaging^{7,8}. However, these techniques generally require the sample to be immediately next to the imaging element. Imaging of buried interfaces in live cells would be prohibitively complicated using such strategies.

By comparison, STED instrumentation is directly compatible with existing approaches

BIOLOGICAL IMAGING

The diffraction barrier broken

Garth J. Simpson

The conventional optical limitations of fluorescence microscopy have been defied, to achieve nanoscale resolution of individual vesicle organelles at the junctions of neuronal cells.

Traditional optical microscopy cannot easily distinguish objects separated by less than about half the wavelength of visible light. For example, measurements of two fluorescent particles closer together than this 'diffraction barrier' generally produce one indistinct, bright blob. In practice, this equates to a maximum resolution corresponding to distances of around 200 nm for biological imaging using visible light. However, Stefan Hell and colleagues have developed a technique that overcomes the diffraction barrier and can resolve objects separated by less than about 40 nm (refs 1–3). On page 935 of this issue, Jahn, Hell and colleagues⁴ use the technique to provide the first *in vitro* images of the movements of single neuronal vesicles — the membrane-bound, bubble-like organelles that mediate communication between neurons. Their images help to settle a question that has been exercising cell biologists for some time.

The latest approach uses stimulated emission depletion (STED) microscopy^{1–3}, a technique that relies on the overlap of two light beams in the focal region. The first beam excites molecules just as in traditional fluorescence microscopy — as the molecule absorbs the energy from the light, it is promoted up to a higher energetic state, and as it relaxes back to the ground state, it releases the energy in the form of light. The second beam, at a different wavelength, suppresses this fluorescence by 'stimulated emission', in which molecules are actively pumped down out of the excited state by light (in essence, absorption driven backwards). In sufficiently intense optical fields, stimulated emission becomes more efficient than fluorescence and

is the dominant effect, drastically reducing the fluorescence.

In STED microscopy, the normal fluorescence is suppressed by a depletion beam that is shaped like a doughnut and contains a node of zero intensity — a hole — at the centre of the focus. So, any fluorescent light that is detected arises just from this hole of about 50 nm across where the depletion beam is absent (Fig. 1). Scanning this narrowed viewing point across the sample allows a higher-resolution image to be built up, much as a computer screen made up of smaller pixels will provide a sharper image (Fig. 1d on page 936 shows the distinct improvement in resolution with this technique). STED microscopy was first demonstrated in one dimension¹ in 1999, and advanced to three dimensions the following year².

Jahn and colleagues⁴ now use this method to investigate vesicle transport in living neurons at single-vesicle resolution. In neurons, vesicles are responsible for transporting neurochemicals to the cell surface, where they fuse with the cell membrane and discharge their contents into the synapses, or gaps, between neurons, thereby transmitting information from one neuron to its neighbours. Following fusion, the vesicle membrane is recycled back inside the cell to form a new vesicle that is refilled with neurotransmitter ready for the next round of communication. The initial fate of the vesicle once it has released its contents is a matter of some contention. One theory — termed 'kiss-and-run' — is that the vesicle membrane remains as a separate entity from the main plasma membrane, like a drop of oil on water, to allow efficient recycling of the membrane-bound molecules and proteins

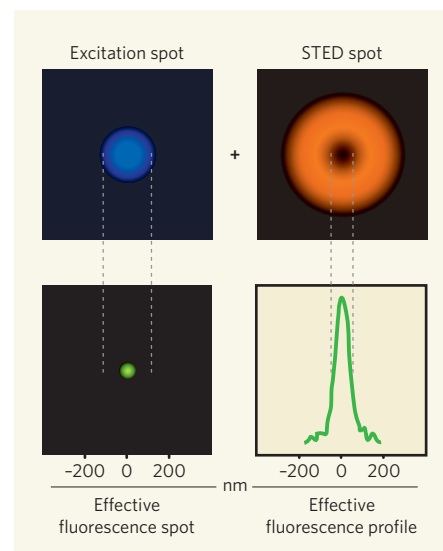


Figure 1 | Stimulated emission depleted (STED) microscopy. The blue excitation beam excites the fluorophore under observation, and is focused down to a diffraction-limited spot. The orange STED beam 'de-excites' the molecules, by pumping them down out of the excited state before they fluoresce. The beam is shaped like a doughnut, with a hole in the centre. By superimposing the STED beam on the excitation beam, the region from which excited fluorophores will fluoresce is reduced to about 50 nm by the STED beam — a much smaller distance than the diffraction limit. Jahn, Hell and colleagues⁴ use this technique to follow vesicles as they fuse with the plasma membrane of living neurons, at a resolution that allows single vesicles to be distinguished.

for optical biological microscopy. Resolution along the z -axis is actually even better than in the x - y plane, making STED well suited for imaging that requires control of the depth of field, such as confocal sectioning, and for three-dimensional reconstructions made by stacking up the x - y sections⁹. Perhaps most significantly, there are no fundamental limits on the spatial resolution that can be achieved by STED. In principle, increasing the intensity of the depletion beam will generally increase the degree of spatial confinement.

The STED approach is not applicable to every molecular system, however. The stimulated emission process usually requires the molecules to fluoresce. This is not a problem in fluorescence microscopy but could be an issue for other applications in which sub-diffraction limited excitation is desirable, such as photolithography. Additionally, the molecules' fluorescence emission profile must be sufficiently different from their excitation profile to ensure minimal absorption and excitation by the STED beam itself.

There is always a risk of sample damage from heating when relatively high-intensity

pulsed lasers are used, but the laser powers required for STED are comparable to those routinely used for multiphoton fluorescence imaging. Furthermore, related approaches³ can be performed at significantly lower laser power, complementing the STED technique. It will be exciting to see the innovations and applications spawned by STED over the next decade — the capabilities of these powerful emerging techniques are only just beginning to be realized. ■

Garth J. Simpson is in the Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA.

e-mail: gsimpson@purdue.edu

1. Klar, T. A. & Hell, S. W. *Opt. Lett.* **24**, 954–956 (1999).
2. Klar, T. A., Jakobs, S., Dyba, M., Egner, A. & Hell, S. W. *Proc. Natl Acad. Sci. USA* **97**, 8206–8210 (2000).
3. Hofmann, M., Eggeling, C., Jakobs, S. & Hell, S. W. *Proc. Natl Acad. Sci. USA* **102**, 17565–17569 (2005).
4. Willig, K. I., Rizzoli, S. O., Westphal, V., Jahn, R. & Hell, S. W. *Nature* **440**, 935–939 (2006).
5. Pohl, D. W., Denk, W. & Lanz, M. *Appl. Phys. Lett.* **44**, 651–653 (1984).
6. Harootunian, A. et al. *Appl. Phys. Lett.* **49**, 674–676 (1986).
7. Smith, D. R. et al. *Science* **305**, 788–792 (2004).
8. Ramakrishna, S. A. *Rep. Prog. Phys.* **68**, 449–521 (2005).
9. Egner, A. & Hell, S. W. *Trends Cell Biol.* **15**, 207–215 (2005).

BEHAVIOURAL ECOLOGY

Grasshoppers don't play possum

Graeme Ruxton

When about to be eaten by frogs, certain grasshoppers assume a static, unwieldy pose that means they cannot be swallowed. Similar behaviours, interpreted as feigning death, may also be open to alternative explanations.

When gripped by a predator, many animals stop struggling and apparently feign death. This behaviour has been described in mammals, birds, fish, reptiles, crustaceans and especially insects. But its evolutionary benefit — or adaptive significance — has generally not been considered, except to speculate that predators somehow 'lose interest' in prey that they mistakenly think are dead.

As reported in *Proceedings of the Royal Society*, Atsushi Honma and colleagues¹ provide a demonstration of an anti-predator benefit to such behaviour. In laboratory experiments, they found that grasshoppers (*Criotettix japonicus*) adopted a characteristic static pose when grabbed by frogs, such that three body parts were stretched in different directions, making the grasshoppers more difficult to swallow. Frogs often abandoned prey that struck this attitude. In further experiments, the grasshoppers did not adopt this pose when captured by other predators that do not swallow them whole. This is both a demonstration of an advantage to predator-induced immobility and a warning that we are probably mistakenly interpreting other behaviours as feigning death.

The adoption of an immobile pose when an



animal is physically restrained is also called tonic immobility, thanatosis and (more colloquially) playing possum. This response is quite different from freezing to make detection by predators more difficult, because it generally happens only after the predator has grasped the prey. In poultry, the duration of tonic immobility induced by human manual restraint is used as a measure of stress². And next time you see a beetle scuttling across the ground, you can attempt to induce tonic immobility your-

self by tapping the beetle lightly on the back.

The usual explanation is that this behaviour fools the predator into thinking that the prey is dead, causing the predator to lose interest. But there is no convincing explanation for this apparently maladaptive loss of interest. Feigning death might be adaptive if the predator is attempting to subdue several prey in rapid succession before returning to consume them, with prey immobility causing the predator to prematurely release one victim and move on to subduing the next³. Even if this explanation is true, however, it cannot explain the impressive diversity of this behaviour throughout the animal kingdom.

Another potential mechanism is that tonic immobility serves as a signal that an animal possesses defensive chemicals. Miyatake *et al.*⁴ have shown that tonic immobility is a heritable trait in *Tribolium* flour beetles, and that it is used to good effect when a beetle is attacked by spiders. They report that, when preying on flies, a spider never set a fly free after grabbing it, and always ate it immediately. In contrast, a spider always relinquished its hold on a flour beetle during the first attack. Miyatake *et al.* found that if the beetle struggled, it was attacked again and often consumed. But they say that, "If, however, the prey feigned death and remained immobile after the first attack, the spider often lost interest and the prey survived".

I suggest that, rather than feigning death, the beetles may be signalling that because they have a strong chemical defence, which the spider may have already sampled in the first attack, they have no need to struggle or flee. This is similar to the characteristically slow flying adopted by butterflies with chemical defence, signalling that they have no need to fear capture⁵. This hypothesis could be tested: I would predict that beetle lines bred for tonic immobility might also be expected to have greater chemical defence than control lines.

So much for speculation. By contrast, Honma *et al.*¹ have produced what is to my knowledge the first logically and empirically supported theory to account for tonic immobility: making physical handling of prey more challenging. It is unlikely that this mechanism can explain more than a subset of cases of tonic immobility. But their paper will do the valuable service of shaking us from the rut of interpreting such behaviours uncritically as feigning death. ■

Graeme Ruxton is in the Division of Environmental and Evolutionary Biology, University of Glasgow, Glasgow G12 8QQ, UK. e-mail: g.ruxton@bio.gla.ac.uk

1. Honma, A. K., Oku, S. & Nishida, T. *Proc. R. Soc. Lond. B* **10**, 1098/rsps.2006.3501 (2006).
2. Campo, J. L., Gil, M. G., Dávila, S. G. & Muñoz, I. *Poultry Sci.* **84**, 1004–1009 (2005).
3. Ruxton, G. D., Sherratt, T. N. & Speed, M. P. *Avoiding Attack: The Evolutionary Ecology of Camouflage, Warning Signals and Mimicry* (Oxford Univ. Press, 2004).
4. Miyatake, T. et al. *Proc. R. Soc. Lond. B* **271**, 2293–2296 (2004).
5. Srygley, R. B. *Proc. R. Soc. Lond. B* **271**, 589–594 (2004).

BRIEF COMMUNICATIONS

A catfish that can strike its prey on land

This denizen of tropical swamps may shed light on how ancient fish were able to survive out of water.

An important step towards understanding the evolution of terrestriality in vertebrates is to identify how the aquatic ancestors of tetrapods were able to access ground-based prey. We have discovered that the 'eel catfish' *Channallabes apus*, an inhabitant of the muddy swamps of tropical Africa, has a remarkable ability to forage and capture prey on land. The animal's capacity to bend its head down towards the ground while feeding seems to be an essential feature that may have enabled fish to make the transition from an aquatic to a terrestrial mode.

The acquisition of a terrestrial lifestyle, one of the key events in the evolution of vertebrates¹, made vastly different physical demands from a life in water². Early tetrapods were confronted with serious constraints on terrestrial foraging because aquatic vertebrates typically suck up their prey, using a rapid rostro-caudal expansion of the mouth cavity^{3,4}. Air is about 800 times less dense than water, so a suction-induced flow will not move prey on land². Early tetrapods were thus forced to develop alternative strategies for capturing their prey.

We discovered that eel catfish, despite being typical suction feeders in water, also capture food on land: the natural diet of this species consists mainly of terrestrial insects⁵. Our observations reveal that, after propelling its head on to the shore, the eel catfish lifts the front part of its body and bends its head downwards. Once the animal makes contact with its prey, the mouth opens and closes repeatedly until the object is held firmly between the jaws (for movie, see supplementary information).

The eel catfish preserves the rostro-caudal expansion wave (mouth opening, followed by extensive hyoid depression) during terrestrial feeding (Fig. 1). The resulting flow of air does not usually transport prey, but it may help when feeding on large, soft items (for movie, see supplementary information). In any case, it seems that prey can be caught on land with relatively minor changes in the higher-level control circuits that guide food capture.

Bending the head ventrally, an action that is also observed in terrestrially feeding mudskippers⁶ (Periophthalmidae), is essential because the prey might otherwise be pushed forwards and because the hyoid can be freely depressed in this position⁷. The posture calls for dorso-ventral flexion along the vertebral column. Recently discovered tetrapod-like fish⁸ and early tetrapods such as *Ichthyostega*⁹

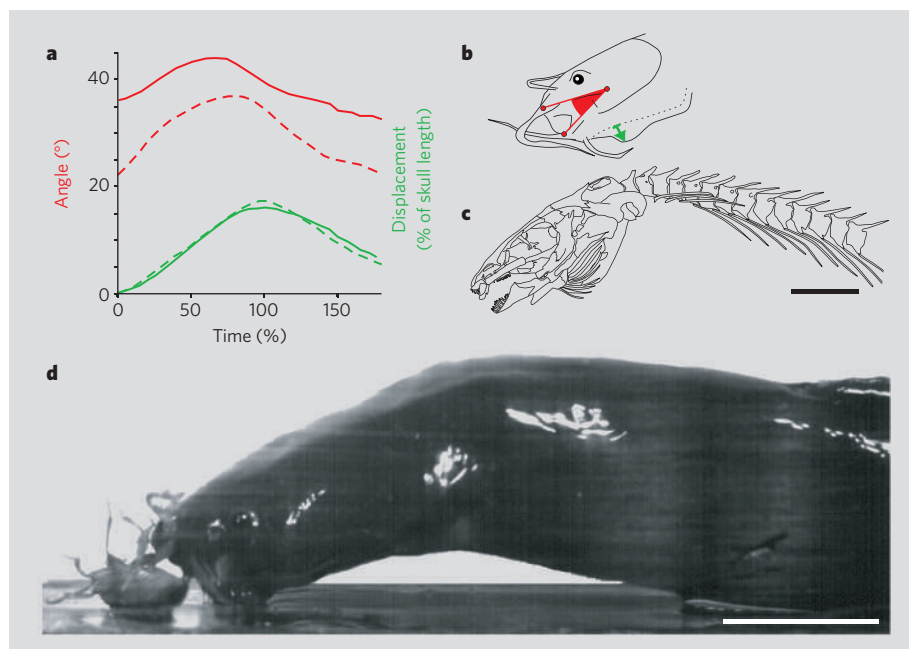


Figure 1 | Prey capture by the eel catfish *Channallabes apus*. **a**, Average, time-scaled kinematic profiles of the gape angle (red) and hyoid depression (green) in catfish capturing land-based prey (solid lines) compared with catfish bottom-feeding in water (dashed lines). Time value of 100% corresponds to maximal hyoid depression (see supplementary information). Note that the timing and range of mouth opening and hyoid depression are similar in both environments, resulting in a rostro-caudal sequence of head expansion. **b**, Measurement of gape angle (red) and hyoid depression (green). **c**, Posture of the skull and anterior part of the vertebral column during terrestrial feeding. **d**, High-speed-video frame of *C. apus* capturing a beetle on land. The inclination of the head towards the substrate is much steeper during terrestrial than during aquatic prey capture. Scale bars, 10 mm.

show adaptations for such dorso-ventral flexion of the presacral vertebral column, and this may have allowed these animals to capture prey on land more effectively.

Our findings also show that having weight-bearing pectoral fins is not a prerequisite for capturing prey on land in a fish that has a flexible body. The importance of dorso-ventral flexion in the initial acquisition of a terrestrial feeding mode is further supported by the reduced number of presacral vertebrae that are found in tetrapod-like fish making the transition to tetrapods with weight-bearing pectoral fins⁸.

Sam Van Wassenbergh*, **Anthony Herrel***, **Dominique Adriaens†**, **Frank Huysentruyt†**, **Stijn Devaere†**, **Peter Aerts***

*Department of Biology, University of Antwerp, 2610 Wilrijk, Belgium

e-mail: sam.vanwassenbergh@ua.ac.be

†Department of Biology, Ghent University, 9000 Ghent, Belgium

‡Department of Movement and Sports Sciences, Ghent University, 9000 Ghent, Belgium

1. Ahlberg, P. E. & Milner, R. *Nature* **368**, 507–514 (1994).
2. Bramble, D. M. & Wake, D. B. in *Functional Vertebrate Morphology* (eds Hildebrand, M. et al.) 230–261 (Belknap, Cambridge, Massachusetts, 1985).
3. Lauder, G. V. *J. Morphol.* **163**, 283–317 (1985).
4. Summers, A. P. et al. *J. Exp. Zool.* **281**, 280–287 (1998).
5. Huysentruyt, F. et al. *Belg. J. Zool.* **134**, 25–30 (2004).
6. Sponder, D. L. & Lauder, G. V. *J. Zool., Lond.* **193**, 517–530 (1981).
7. Van Wassenbergh, S. et al. *J. Exp. Biol.* **208**, 4627–4638 (2005).
8. Daeschler, E. B., Shubin, N. H. & Jenkins, F. A. Jr *Nature* **441**, 757–763 (2006).
9. Ahlberg, P. E. et al. *Nature* **437**, 137–140 (2005).

Supplementary information accompanies this communication on Nature's website.

Received 27 January; accepted 17 March 2006.

Competing financial interests: declared none.

doi:10.1038/440881a

BRIEF COMMUNICATIONS ARISING online
 www.nature.com/bca see Nature contents.

TAXONOMY

Sus bucculentus revisitedArising from: C. P. Groves, G. B. Schaller, G. Amato & K. Khounboline *Nature* **386**, 335 (1997)

In 1997, the rediscovery of *Sus bucculentus* in Laos was announced by Groves *et al.*¹ — this wild pig species had gone unrecorded since first being described in 1892. Although the identification of the new specimen was based initially on morphology, the authors also used a 7% sequence divergence from the common Eurasian pig *S. scrofa* (based on their analysis of 327 base pairs of the gene encoding mitochondrial 12S ribosomal RNA) as support for the species status of *S. bucculentus*. Concerned about the large divergence reported for a relatively conserved gene, and the absence of the sequence in any public database, we analysed an additional tissue sample from the specimen and found only 0.6% divergence from *S. scrofa*. Our more extensive analysis places the sample within the *S. scrofa* clade, calling into question the species status of *S. bucculentus* and demonstrating the need for both phylogenetic and morphological evidence in defining species.

We extracted DNA from a dried tissue sample of the more recently described *S. bucculentus* specimen¹ (kindly supplied by C. P. Groves). The D-loop, 12S and cytochrome *b* regions of mitochondrial DNA were amplified, sequenced and compared with a wide range of sequences from GenBank and from our laboratory database. The comparison included three samples of *S. scrofa* from archaeological sites on Pacific islands; these were extracted in a dedicated ancient-DNA laboratory using standard precautions to avoid contamination², and were independently replicated by J. Hay (Massey University, New Zealand) in an ancient-DNA laboratory in which porcine DNA had not previously been processed.

Phylogenetic analysis of the 12S-rRNA gene would be uninformative because of poor species representation in GenBank, so we used a barcode approach³ to compare intraspecific and interspecific sequence distances⁴. For cytochrome *b* and D-loop, phylogenetic trees were estimated using both bayesian⁵ and maximum-likelihood⁶ methods.

For the 12S-rRNA gene, the mean pairwise distance between *S. bucculentus* and *S. scrofa* (0.006) is identical to that within *S. scrofa* (0.006), and only a tenth of the pairwise distance (0.055) between *S. scrofa* and the African warthog (*Phacochoerus africanus*). Unlike the 7% sequence difference reported earlier¹, this 0.6% difference is consistent with the conserved nature of the 12S-rRNA gene. For cytochrome *b* and D-loop, all four phylogenetic trees have similar topology, placing *S. bucculentus* within the Asian subclade of *S. scrofa*. Furthermore, in the D-loop analyses (Fig. 1), *S. bucculentus* sits in a clade that

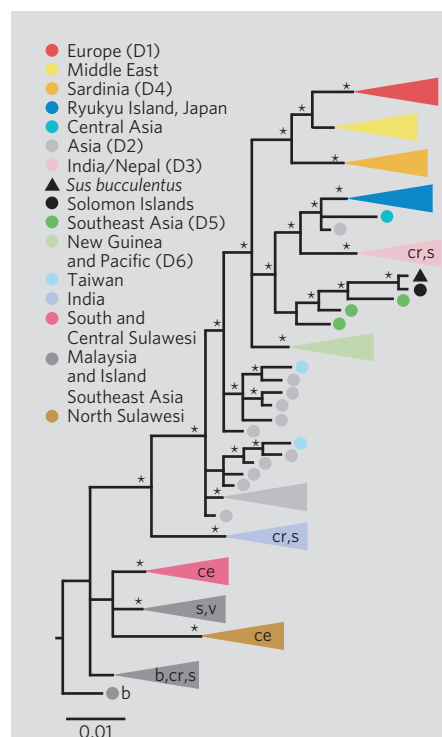


Figure 1 | Simplified phylogenetic tree based on bayesian analysis of mitochondrial D-loop region DNA for several porcine species (*Sus* spp.), rooted using the African warthog (not shown). Triangles represent monophyletic clades and circles represent paraphyletic groups and single sequences. Clades are coloured and named according to ref. 9. Nodes marked by asterisks have posterior $P < 80\%$. All clades represent *S. scrofa* except when labelled: b, *S. barbatus*; ce, *S. celebensis*; cr, *S. cristatus*; s, *S. scrofa*; v, *S. verrucosus*. GenBank accession numbers for the new sequences used in this study are DQ444703–DQ444711. Sample details, methods of DNA extraction, amplification and sequencing, computational methods, the full phylogenetic trees based on the D-loop and cytochrome *b* regions of the mitochondrial DNA, and the barcoding analysis of the 12S gene are available from the authors (www.cebl.auckland.ac.nz/~hros001/Susbucc/).

includes *S. scrofa* from Thailand and Myanmar, countries that border Laos. Coincidentally, an archaeological specimen from Tikopia in the Solomon Islands, albeit from an uncertain chronological context⁷, had a D-loop sequence identical to that of *S. bucculentus*.

These results undermine the species status of *S. bucculentus*. The partial skull from 1997 was morphologically identical to the 1892 skull¹. But Groves and colleagues have recently judged the key taxonomic characters of the mandibular canine and malar tuberosity to be less distinctive than once thought, even though multivariate analyses of cranial and dental measurements tend to separate the two skulls from samples of Indochinese *S. scrofa*⁸. Our phylogenetic results are consistent with a recently shared ancestry between the 1997 *S. bucculentus* sample from the Annamite Range and Asian domestic and wild *S. scrofa*. Also, our *S. bucculentus* 12S-rRNA gene sequence was within the range of intraspecific variation shown by *S. scrofa*.

There are three possible explanations for these findings, a view shared by C. P. Groves (personal communication). First, hybridization may have introduced a locally occurring mitochondrial haplotype from *S. scrofa* into *S. bucculentus*. Second, speciation and morphological differentiation of *S. bucculentus* may have occurred relatively recently and so subsequent mitochondrial lineage sorting

remains incomplete. A third possibility, however, is that the morphology previously^{1,8} designated as that of a separate species, *S. bucculentus*, is no more than one extreme of the natural range of variation shown by the wild Indochinese *S. scrofa*.

Judith H. Robins*, Howard A. Ross†, Melinda S. Allen‡, Elizabeth Matisoo-Smith*

*Department of Anthropology and Allan Wilson Centre for Molecular Ecology and Evolution, †Bioinformatics Institute and School of Biological Sciences, ‡Department of Anthropology, The University of Auckland, Private Bag 92019, Auckland, New Zealand
e-mail: j.robins@auckland.ac.nz

- Groves, C. P., Schaller, G. B., Amato, G. & Khounboline, K. *Nature* **386**, 335 (1997).
- Handt, O., Höss, M., Krings, M. & Pääbo, S. *Experientia* **50**, 524–529 (1994).
- Hebert, P. D. N., Cywinski, A., Ball, S. L. & deWaard, J. R. *Proc. R. Soc. Lond. B* **270**, 313–321 (2003).
- Swofford, D. L. *PAUP: Phylogenetic Analysis Using Parsimony (*and Other Methods)* (Sinauer, Sunderland, Massachusetts, 2002).
- Huelsenbeck, J. P. & Ronquist, F. *Bioinformatics* **17**, 754–755 (2001).
- Guindon, S. & Gascuel, O. *Syst. Biol.* **52**, 696–704 (2003).
- Kirch, P. V. & Yen, D. E. in *Tikopia: The Prehistory and Ecology of a Polynesian Outlier* 100 (Bishop Museum, Honolulu, 1982).
- Groves, C. P. & Schaller, G. B. in *Antelopes, Deer, and Relatives* (eds Vrba, E. S. & Schaller, G. B.) 261–282 (Yale Univ. Press, New Haven and London, 2000).
- Larson, G. *et al. Science* **307**, 1618–1621 (2005).

doi:10.1038/nature04770

Asa Issie, Aramis and the origin of *Australopithecus*

Tim D. White^{1,2}, Giday WoldeGabriel³, Berhane Asfaw⁴, Stan Ambrose⁵, Yonas Beyene⁶, Raymond L. Bernor⁷, Jean-Renaud Boisserie^{1,2,8}, Brian Currie⁹, Henry Gilbert¹, Yohannes Haile-Selassie¹⁰, William K. Hart⁹, Leslea J. Hlusko², F. Clark Howell¹, Reiko T. Kono¹¹, Thomas Lehmann¹², Antoine Louchart¹³, C. Owen Lovejoy¹⁴, Paul R. Renne¹⁵, Haruo Saegusa¹⁶, Elisabeth S. Vrba¹⁷, Hank Wesselman¹⁸ & Gen Suwa¹⁹

The origin of *Australopithecus*, the genus widely interpreted as ancestral to *Homo*, is a central problem in human evolutionary studies. *Australopithecus* species differ markedly from extant African apes and candidate ancestral hominids such as *Ardipithecus*, *Orrorin* and *Sahelanthropus*. The earliest described *Australopithecus* species is *Au. anamensis*, the probable chronospecies ancestor of *Au. afarensis*. Here we describe newly discovered fossils from the Middle Awash study area that extend the known *Au. anamensis* range into northeastern Ethiopia. The new fossils are from chronometrically controlled stratigraphic sequences and date to about 4.1–4.2 million years ago. They include diagnostic craniodental remains, the largest hominid canine yet recovered, and the earliest *Australopithecus* femur. These new fossils are sampled from a woodland context. Temporal and anatomical intermediacy between *Ar. ramidus* and *Au. afarensis* suggest a relatively rapid shift from *Ardipithecus* to *Australopithecus* in this region of Africa, involving either replacement or accelerated phyletic evolution.

The last quarter-century of research into hominid evolution in Africa greatly extended knowledge of early *Australopithecus*. Discoveries at Hadar and Laetoli during the 1970s led to the recognition of *Au. afarensis*, described as a geographically and ecologically widespread, bipedal, megadont, small-brained hominid species lineage. Until recently, the origins of *Australopithecus* were obscured by a sparse fossil record¹.

In 1994, the smaller-toothed, more primitive hominid *Ar. ramidus* was described from Aramis, Ethiopia² (hominid refers to the human clade subsequent to divergence from our common ancestor with chimpanzees^{3,4}). These finds were followed in rapid succession by the description of *Au. anamensis* from Kenya^{5–8} and the naming of three Late Miocene taxa (*Ardipithecus kadabba*, Ethiopia, ~5.5–5.8 million years (Myr) ago^{3,9}; *Orrorin tugenensis*, Kenya, ~5.7–6.0 Myr ago^{10,11}; and *Sahelanthropus tchadensis*, Chad, ~6–7 Myr ago¹²). Relative to extant and extinct apes, these taxa display derived craniodental and post-cranial characters suggesting that they are all cladistically hominid. Their phylogenetic relationships and locomotor capabilities are under active study and debate^{13,14}.

In contrast, the time-successive species *Au. anamensis* and *Au. afarensis* are widely interpreted as sampling an evolving lineage^{4–8}. All known *Au. anamensis* specimens date to between ~3.9 and ~4.2 Myr ago. The earliest definitive *Au. afarensis* is at Laetoli,

~3.6 Myr ago¹⁵. This younger chronospecies is known by partial skeletons, well-preserved skulls and even attributed footprints. In contrast, *Au. anamensis* was heretofore documented only from the Turkana basin, and represented there by a relatively small sample⁸.

We report here on newly recovered Pliocene fossils from the Middle Awash study area, Afar rift, Ethiopia. The Adgantole Member of the Sagantole Formation¹⁶ has yielded a hominid maxilla. Contemporary sediments exposed approximately 10 km to the west have yielded an additional 30 hominid specimens representing a minimum of eight individuals (Figs 1–3; see also Supplementary Table 1). Dated to ~4.12 Myr ago and attributed to *Au. anamensis*, these remains extend the known range of this taxon by about 1,000 km to the northeast and extend the anatomical representation of early *Australopithecus*.

Geology, geochronology and palaeoenvironment

The Central Awash Complex of the Middle Awash study area includes the well-dated >300-m-thick Sagantole Formation¹⁶. Vertebrate fossil assemblages containing *Ar. ramidus* are known from nearby Lower Aramis Member strata of this formation, dated by ⁴⁰Ar–³⁹Ar to 4.416 ± 0.031 and 4.419 ± 0.068 Myr ago (previous ⁴⁰Ar–³⁹Ar ages¹⁶ are recalculated herein to reflect a revised age for the Fish Canyon sanidine standard¹⁷). Aramis locality 14 is located stratigraphically

¹Human Evolution Research Center, Museum of Vertebrate Zoology, 3101 Valley Life Sciences Building, and ²Department of Integrative Biology, University of California at Berkeley, Berkeley, California 94720, USA. ³Hydrology, Geochemistry and Geology Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA. ⁴Rift Valley Research Service, P.O. Box 5717, Addis Ababa, Ethiopia. ⁵Department of Anthropology, University of Illinois, Urbana, Illinois 61801, USA. ⁶Department of Anthropology and Archaeology, Authority for Research and Conservation of the Cultural Heritage, Ministry of Youth, Sports and Culture, P.O. Box 6686, Addis Ababa, Ethiopia. ⁷College of Medicine, Department of Anatomy, Laboratory of Evolutionary Biology, Howard University, Washington DC 20059, USA. ⁸Département Histoire de la Terre, USM 0203, UMR 5143 CNRS, Unité Paléobiodiversité et Paléoenvironnement, Muséum National d'Histoire Naturelle, Paris 75005, and Laboratoire de Géobiologie, Biochronologie et Paléontologie Humaine, UMR 6046, Université de Poitiers, 86022 Poitiers Cedex, France. ⁹Department of Geology, Miami University, Oxford, Ohio 45056, USA. ¹⁰Department of Physical Anthropology, Cleveland Museum of Natural History, 1 Wade Oval Drive, Cleveland, Ohio 44106, USA. ¹¹Department of Anthropology, National Science Museum, Hyakunincho, Shinjuku-ku, Tokyo 169-0073, Japan. ¹²Palaeontology Section, Transvaal Museum, P.O. Box 413, Pretoria, South Africa. ¹³Laboratoire Paléoenvironnements et Paléobiosphère, UMR 5125, Université Claude Bernard, Lyon 1, 69622 Villeurbanne Cedex, France. ¹⁴Department of Anthropology and Division of Biomedical Sciences, Kent State University, Kent, Ohio 44242, USA. ¹⁵Berkeley Geochronology Center, 2455 Ridge Road, Berkeley, California 94709, and Department of Earth and Planetary Science, University of California at Berkeley, Berkeley, California 94720, USA. ¹⁶Institute of Natural and Environmental Sciences, University of Hyogo, Yayoigaoka, Sanda 669-1546, Japan. ¹⁷Department of Geology and Geophysics, Yale University, New Haven, Connecticut 06520, USA. ¹⁸Natural History Museum, Sierra College, Rocklin, California 95677, USA. ¹⁹The University Museum, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

~80 m above the *Ar. ramidus*-bearing interval, in the overlying Adgantole Member. It has yielded a single hominid maxilla. The Asa Issie area (ASI) lies ~10 km west of Aramis locality 14, separated from the Central Awash Complex by a major northwest–southeast trending fault (Fig. 1). Three localities in this area (ASI-VP-2 and ASI-VP-5, and HAN-VP-1 at Hana Hari) have yielded additional hominid fossils from sediments contemporaneous with the Saganatole Formation's Adgantole Member.

The age estimate of ~4.1–4.2 Myr ago for Aramis 14 and ASI-VP-2 and ASI-VP-5 (see Methods) is confirmed by the presence of

biochronologically sensitive taxa. Compared to Aramis Member precursors, *Anancus*, aff. *Hippopotamus*, *Nyanzachoerus jaegeri* and *Kuseracolobus*¹⁸ are all more evolutionarily derived and consistent with an age younger than 4.4 Myr (Gàala tuff)¹⁶ but older than 3.9 Myr ago (Moiti tuff)¹⁶.

The Aramis locality 14 maxilla was unaccompanied by substantial associated faunal remains. In contrast, the hominids from ASI-VP-2 and ASI-VP-5 are stratigraphically and spatially tightly associated with over 500 vertebrate fossils for which assemblage composition and relative abundance provide strong, high-fidelity palaeoenvironmental

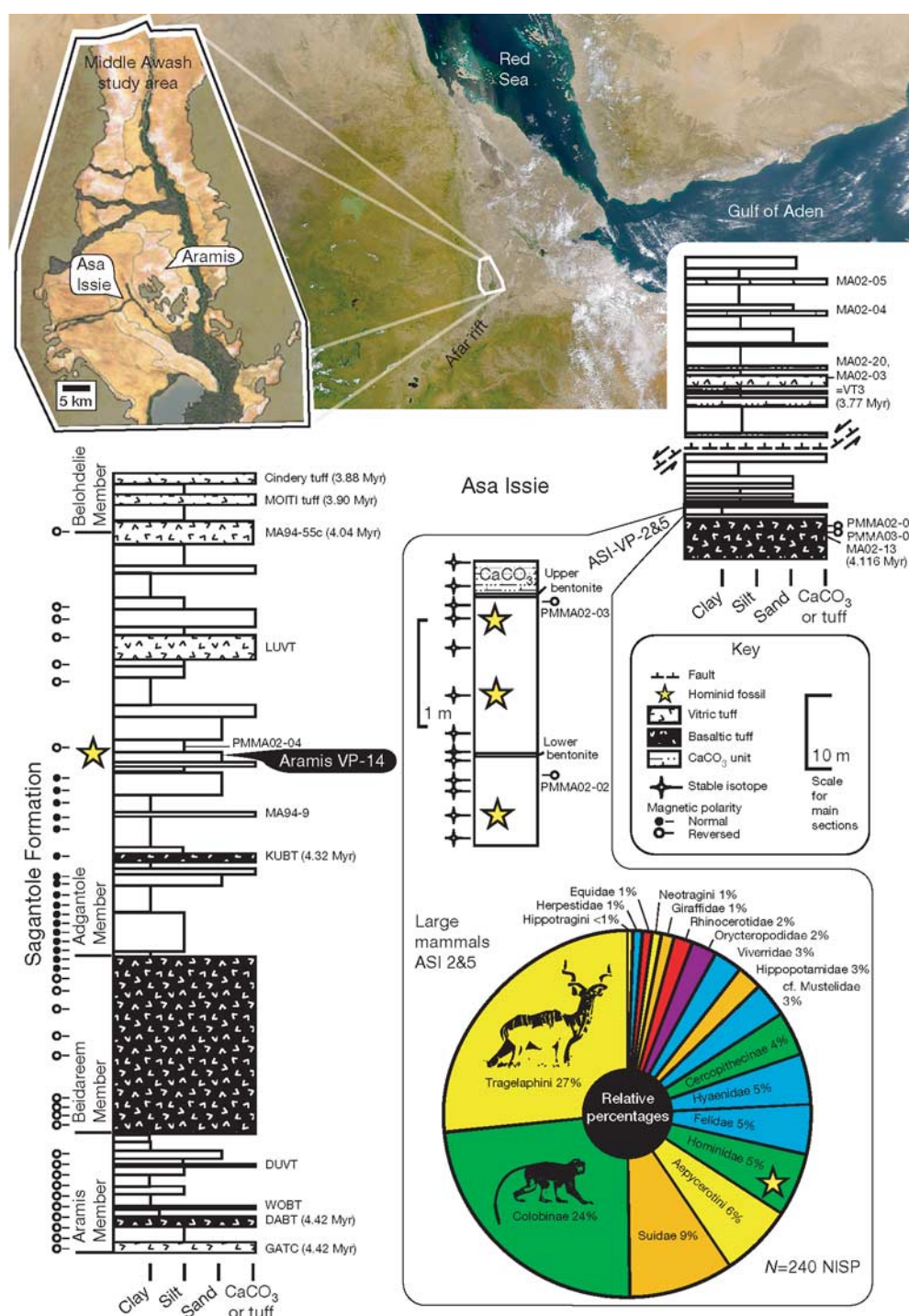


Figure 1 | Geography, stratigraphy, chronology and faunal background for the Asa Issie hominids. The chart at the bottom right shows relative abundance of the taxa indicated, based on NISP values reported in

Supplementary Table 3. DABT, Daam Aatu basaltic tuff; DUVT, Dummu vitric tuff; GATC, Gàala tuff complex; LUVT, Lubaka vitric tuff; PMMA, Middle Awash palaeomagnetic sample; WOBT, Wodara basaltic tuff.

signals. Isotopic analysis of palaeosols interbedded with these vertebrate fossils (Fig. 1) provides average carbonate root cast and nodule $\delta^{13}\text{C}_{\text{PDB}}$ and $\delta^{18}\text{O}_{\text{SMOW}}$ values (where PDB is Pee Dee belemnite and SMOW is standard mean ocean water) that reflect humid, grassy, woodland savannah environments ($\sim 25\text{--}35\%$ C_4 grass; Supplementary Table 2 and Supplementary Fig. 1). This is isotopically intermediate between the generally more closed, cooler and/or humid woodland habitats of earlier Mio-Pliocene and open, warmer and drier later Pliocene and early Pleistocene wooded grassland environments represented by eastern African Rift Valley palaeosol carbonate records^{19–23}.

The ASI-VP-2 and ASI-VP-5 sediments are interpreted as having accumulated on a flood plain distal to the main channel. Silts and clays with interbedded nodular palaeosol horizons yielded a fragmented vertebrate fauna with no evidence of fluvial and/or out-of-habitat transport²⁴. Sedimentology and the combined faunal assemblage indicate strong similarity in taphonomic history with the vertebrate assemblage at the older *Ar. ramidus* ARA-VP-1 and ARA-VP-6 localities (heavy carnivore ravaging followed by rapid burial with little or no transport)².

The ASI-VP-2 and ASI-VP-5 faunal assemblage is notable for the rarity or absence of aquatic elements (fish, crocodile, hippopotamus, waterfowl, freshwater gastropods). Rather, the assemblage appears to have been primarily terrestrially emplaced. Primates and bovids predominate, with the former most abundant (221 and 113, respectively, of 540 total identifiable specimens; Fig. 1). Colobine primates outnumber cercopithecines by 57:9. Alcelaphine and reduncine bovids are absent, and *Tragelaphus* outnumbers all other bovids assignable to tribe at a ratio of 64:18 identifiable dental specimens. The avifauna and micromammalian fauna yield taxonomic profiles, abundance values and surface modifications that closely parallel

those of the *Ar. ramidus*-associated fauna from Aramis, with *Atherurus*, *Oenomys* and *Taphozous* (the forest species) reflective of heavily wooded habitats. From these faunal associations it is evident that the Asa Issie hominids were closely and regularly associated with a narrow range of habitats varying from closed to grassy woodlands. This is similar to the habitat inferred for *Ar. ramidus* at nearby Aramis ~ 0.2 Myr earlier.

Hominid fossils

Specimen ARA-VP-14/1 is a left maxilla with fragmentary crowns of I^2 and $\text{M}^2\text{--}\text{M}^3$, broken canine, premolar and molar roots, and adjacent palatal and lateral maxillary surface. The right maxilla contains the broken P^4 root and damaged molar crowns. Tooth wear is advanced, with the incisor crown worn to root and a large, deep M^3 protocone dentine exposure. The palate is very shallow anteriorly on the left. Its roof is distorted superiorly on the right. The canine jugum would have formed the margin of the pyriform aperture. The specimen is slightly smaller but anatomically similar in preserved parts to the KNM-KP (Kenya National Museums, Kanapoi site) 29283 *Au. anamensis* paratype (Fig. 2). The canine root is relatively vertically implanted as in KNM-KP 29283 and some *Au. afarensis*. Tooth rows are straight but dental arcade shape can only be approximated.

Associated dental rows ASI-VP-2/2 and ASI-VP-2/334 are from separate individuals. They definitively place the Asa Issie sample within expected ranges of *Au. anamensis* variation. Molar crown dimensions are at or slightly above (ASI-VP-2/334) the upper end of the known *Au. anamensis* range. Combined with the slightly smaller ARA-VP-14/1 dentition, these Middle Awash post-canine teeth are distinctly larger than *Ar. ramidus* but broadly equivalent to both *Au. anamensis* and *Au. afarensis* counterparts.

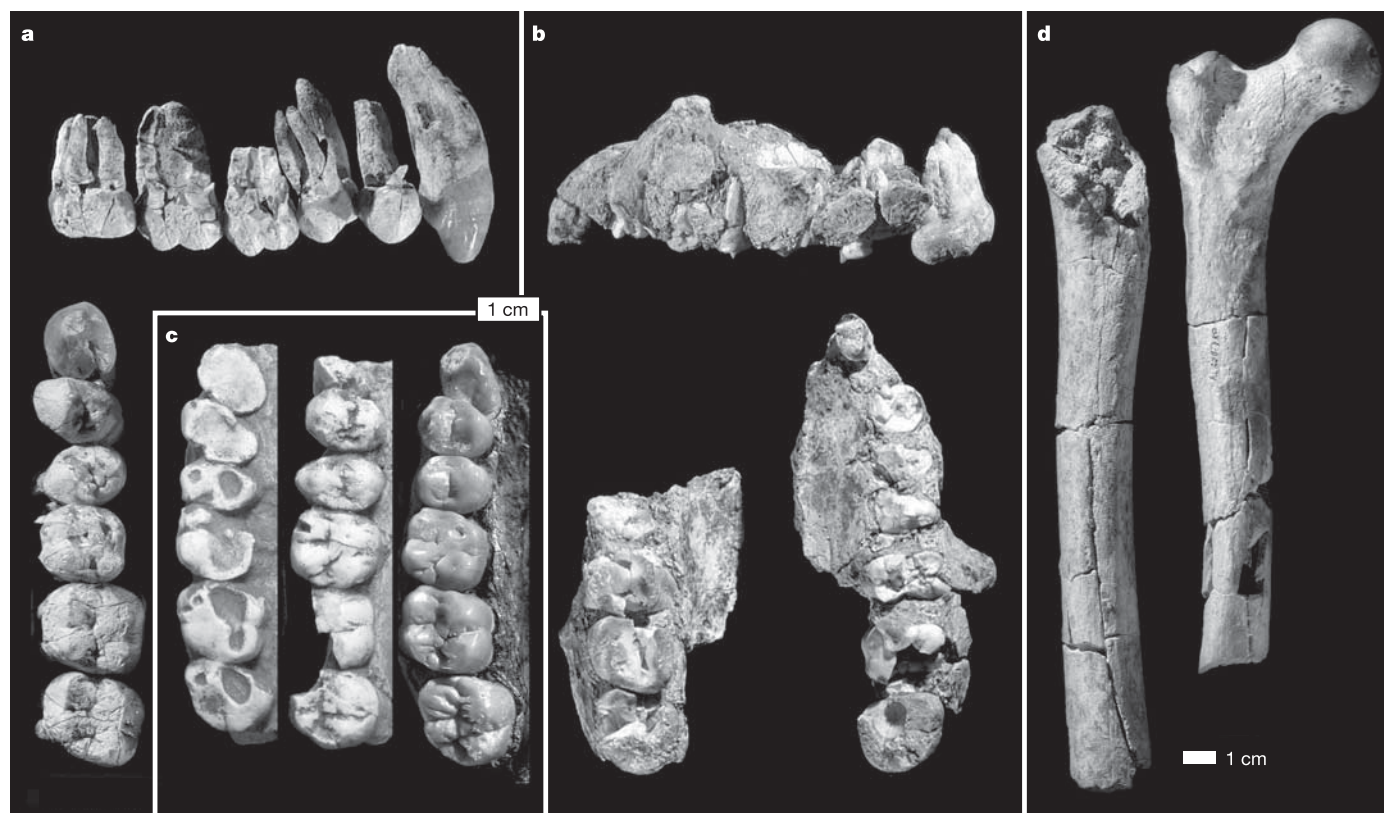


Figure 2 | Aramis and Asa Issie fossil hominids. **a**, ASI-VP-2/334 right maxillary dentition. **b**, ARA-VP-14/1 maxilla with dentition. Alignment of right and left maxillary arcs is approximate. **c**, *Au. anamensis* (KNM-KP 29283 and KNM-ER 30745, left and middle, respectively; casts, reversed)

and *Au. afarensis* (A.L. 200-1, right) dentitions. **d**, Comparison of the ASI-VP-5/154 right femoral shaft with the smaller but otherwise morphologically similar left proximal femur of A.L. 288-1 (Lucy; *Au. afarensis*).

Both ASI-VP-2/2 and ASI-VP-2/334 preserve the diagnostically important upper canines. In absolute crown dimensions, these are large, at or above the known *Au. anamensis* and *Au. afarensis* ranges of variation; isolated canine ASI-VP-2/367 is smaller. The three canines encompass the known Kenyan *Au. anamensis* size range. Asa Issie canine size relative to molar size is comparable to or slightly greater than the two known examples of *Au. anamensis* (KNM-KP 29283, KNM-KP 30498) and intermediate between the *Ar. ramidus*

and *Au. afarensis* conditions (Supplementary Discussion 1). Canine shape (mesiodistal versus buccolingual dimensions) is also intermediate between *Ar. ramidus* and known *Au. anamensis* conditions, tending towards the more mesiodistally elongate morphology considered to be distinctive of *Au. anamensis* but not *Au. afarensis* (Fig. 3; see also Supplementary Fig. 2)⁸. Another feature that we interpret to be of evolutionary significance is the development of the upper canine's mesiolingual ridge. The known Kenyan *Au. anamensis* upper

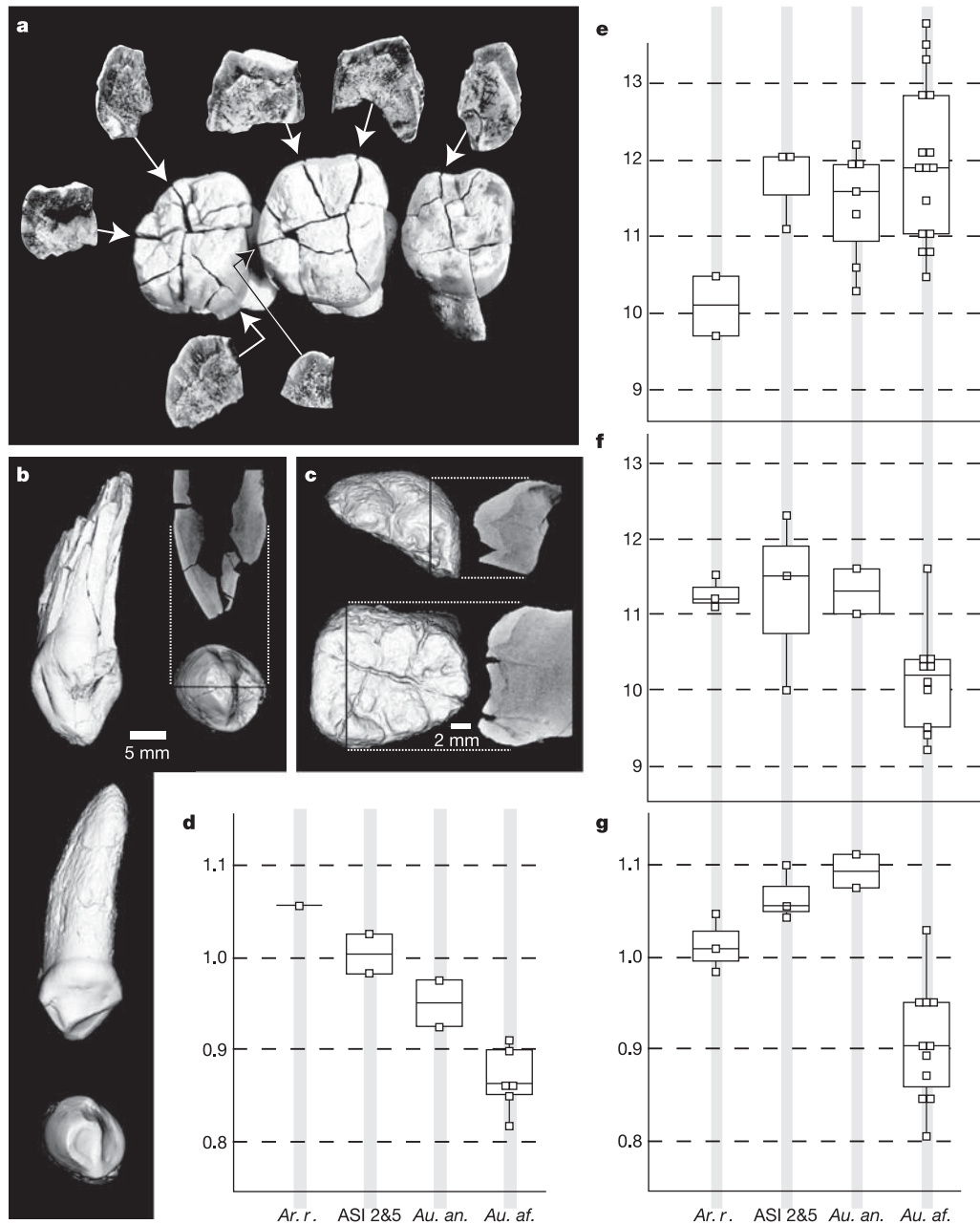


Figure 3 | Dental features of the Asa Issie hominid dentition. **a**, ASI-VP-2/334 maxillary molars (unglued) showing enamel thickness in natural cross-sections. **b**, ASI-VP-2/334 (top) and ASI-VP-2/2 (bottom) maxillary canines, micro-CT-based surface-rendered images and cross-section through the cusp tip. Lingual and occlusal views are shown; note the enamel thickening towards the cusp tip in micro-CT section view (maximal radial thickness is 1.54 mm; Supplementary Discussion 2). Note the mesiodistally elongate crown of ASI-VP-2/2. **c**, ASI-VP-2/146 (top) and ASI-VP-5/1 (bottom) molars, micro-CT-based surface-rendered images and cross-sections. Occlusal and section views are shown. Note the strongly concave buccal crown enamel–dentine junction contour resulting in exaggerated molar

crown flare. **d–g**, Dental metric comparisons among early hominid taxa. **d**, Relative canine size (upper canine maximum dimension divided by upper M¹ mesiodistal length). **e**, Upper M¹ mesiodistal length (in mm). **f**, Upper canine mesiodistal length (in mm). **g**, Upper canine shape (mesiodistal length divided by buccolingual breadth). Each small square represents one specimen; vertical lines are total ranges, whereas horizontal lines are medians and quartiles. *Ar. r.*, *Ar. ramidus*; ASI 2&5, and Aramis locality 14 *Au. anamensis*; *Au. an.*, Kenyan *Au. anamensis*; *Au. af.*, *Au. afarensis*. Comparative sample sizes and sources are outlined in Supplementary Discussion 1.

canines were noted to have a stronger mesiolingual ridge than *Au. afarensis* homologues⁸. *Au. afarensis* upper canines tend to have a more spatulate or incisiform lingual fossa, although expression of the mesiolingual ridge is variable. All three Asa Issie canines show a strong mesiolingual ridge as in KNM-KP 35839 (and slightly less so in KNM-KP 30498). The same ridge is even stronger in known *Ar. ramidus* examples² (ARA-VP-1/300, ARA-VP-6/1).

The upper P³ is asymmetric in shape, as in known *Au. anamensis* and some *Au. afarensis* (Laetoli) examples. The Asa Issie molars tend to be low crowned with flaring buccal and lingual crown faces. The second molar is much larger than the first in ASI-VP-2/334 and less so in the other two available specimens. The Asa Issie lower third molars exhibit a developed distal crown as is commonly the case in *Au. afarensis* and *Au. anamensis* homologues, but not in known examples of *Ar. ramidus* (for example, ARA-VP-1/128)².

Enamel thickness was examined non-invasively either at appropriately broken natural fracture locations (ASI-VP-2/334) or by high-resolution micro-computerized-tomography (micro-CT) visualization²⁵ (Fig. 3; see also Supplementary Discussion 2). Maximum radial thickness of the lateral crown face of the two posterior molars²⁶ ranged between ≥ 1.7 mm to ≥ 2.3 mm in the 'functional-side' cusps (buccal in lowers and lingual in uppers) and ≥ 1.3 mm to 2.0 mm in the opposite-side cusps. This is comparable to known *Au. anamensis* where molar enamel thickness appears to be close to the *Au. afarensis* condition in the functional-side cusps⁸, but varies towards the thinner distribution in the opposite-side cusps. Enamel thickness measures of *Ar. ramidus* have been reported to occupy a thinner range of variation², although considerable within-species variation is currently being documented in modern human and ape control samples^{25,26}.

The ASI-VP-2 and ASI-VP-5 post-crania include a metatarsal shaft without ends, an eroded distal foot phalanx, and an intact intermediate hand phalanx. The last shows slight dorsal longitudinal curvature, accentuated distally. The proximal half of the palmar surface shows deeply excavated attachment sites for m. flexor digitorum superficialis encroaching on a prominent, raised central ridge. The specimen is morphologically similar to those from Hadar, but is longer relative to its breadth. Four vertebral fragments include

an atlas larger than its single Hadar homologue and a thoracic arch larger than any in the Hadar A.L. 288-1 specimen.

One study⁷ predicted that when found, the *Au. anamensis* femur would be similar to that of *Au. afarensis*. Specimen ASI-VP-5/154 is approximately 75% of an adult right femur shaft preserving the base of the lesser trochanter and part of the neck-shaft junction. The shaft is well preserved except for its entire distal-most portion, lost just proximal to the popliteal surface. The shaft retains surface detail but is broken into slightly offset fragments that artificially accentuate the very slight (original) anteroposterior shaft curvature (in its original condition the shaft would have been much straighter). The shaft is remarkable for its thick cortex revealed throughout its length by the broken cross-sections.

A strongly roughened, >3-cm long (superioinferiorly), posterolaterally positioned attachment for m. gluteus maximus represents the most rugose part of the bone and contrasts sharply with the otherwise minimal relief of its shaft. There is no linea aspera, but only relatively blunt outlines of the adductor attachments both medially and laterally. At the shaft's approximate midpoint, these two minimal ridges are separated by about 11 mm, a distance of considerable breadth given the probable original length of the bone. The Asa Issie femur is thereby similar to the 'minimal linea aspera' morphology of the posterolateral femur that characterizes the smaller A.L. 288-1 femur (Fig. 3). In this sense, the older Asa Issie specimen is on the presumably primitive end of the considerable range of variation in *Au. afarensis* with respect to this character.

Early *Australopithecus* environment and biogeography

Palaeoenvironmental circumstances surrounding *Au. anamensis* ~1,000 km to the south in Kenya have been described for Allia Bay as a mixed assemblage sampling aquatic, forest, grassland and bushland^{27,28}. Nearby Kanapoi conspecifics were found in another mix of environments described as dry, possibly open, wooded, or bushland conditions with a wide gallery forest in the vicinity⁵. Habitat preferences in such mixed assemblages are difficult to ascertain despite the assertion²³ that hominids "favored mosaic settings". In contrast, the Ethiopian occurrence of *Au. anamensis* described here allows its tight spatial and temporal placement in a vertebrate assemblage with

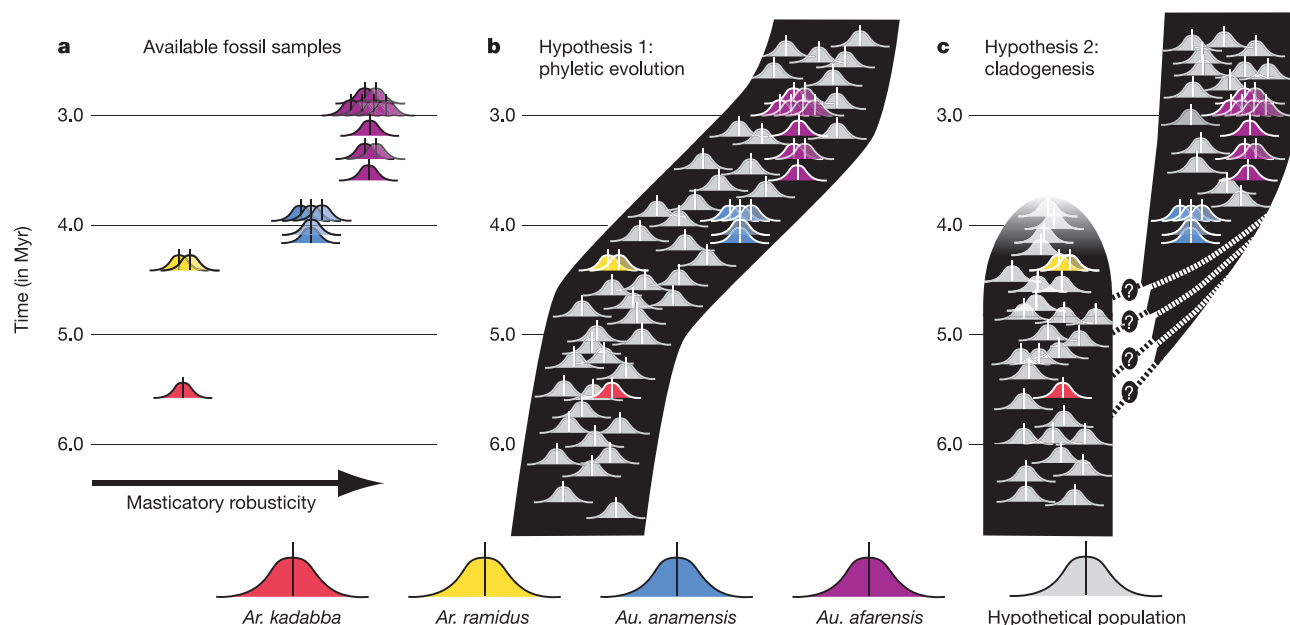


Figure 4 | Phylogenetic hypotheses. Known fossil hominid samples are depicted in colour, by site (for example, Middle Awash for *Ar. kadabba*; Aramis and Gona for *Ar. ramidus*; Kenyan, Tanzanian and Ethiopian occurrences for *Australopithecus*). Currently available samples may be

hypothesized to represent a single lineage evolving at varying rates (phyletic evolutionary origin of *Australopithecus*) or a speciation event (cladistic evolutionary origin of *Australopithecus*). Neither hypothesis can be falsified with available sample densities.

taphonomic integrity. Its relative abundance suggests that it was a regular occupant of a wooded biome that appears to have persisted in this part of the Afar during the 200,000-yr interval subsequent to *Ar. ramidus* at Aramis.

At Aramis, the lone hominoid and largest primate was *Ar. ramidus* (109 of 6,156 identified specimens so far). No trace of *Australopithecus* has been recovered in this (4.4 Myr ago) or any contemporary or older African deposit²⁹ (*contra* refs 30, 31). Furthermore, *Ardipithecus* has not been found at Asa Issie (~4.1–4.2 Myr ago) or in any other contemporary or younger fauna. Thus, *Ardipithecus* and *Australopithecus* are, so far as is known, mutually exclusive in temporal distribution. Defining the first appearance datum of *Australopithecus* is hazardous given the incompleteness of the geological record, but its first appearance in the Turkana basin at three separate sites (Kanapoi, Allia Bay and Fejej) is coincident (within age constraints) with its appearance in the Afar Rift at Aramis, Asa Issie and possibly Galili³².

Phylogenetics of early *Australopithecus*

In an assessment of fossils from Kanapoi (3.9–4.2 Myr ago), the anagenetic series *Ar. ramidus*, *Au. anamensis* and *Au. afarensis* has been hypothesized^{7,8}. The evidence reported here from the Afar Rift constitutes a strong test from a single stratigraphic succession that fails to falsify this hypothesis. Middle Awash *Au. anamensis* is anatomically intermediate in many characters between the earlier *Ar. ramidus* and the later *Au. afarensis* from the same study area (see Supplementary Discussion 3 for cladistic analysis). Twenty years ago, *Au. afarensis* was heralded as the “epitome of australopithecine primitiveness”³³ and a “locomotor missing link”³⁴, but it is decidedly derived relative to *Au. anamensis* and *Ar. ramidus*.

Two phylogenetic hypotheses concerning the origin of *Australopithecus* can be offered to account for the available data. The first hypothesis (Fig. 4) derives *Au. anamensis* phyletically from *Ar. ramidus* within a 200,000-yr interval. The second involves cladogenesis of *Au. anamensis* from an ancestor (presumably *Ardipithecus* or some close relative) even deeper in the Pliocene or Late Miocene. Under the latter hypothesis, *Ar. ramidus* would represent a relict species in an ecological refugium.

Early hominid evolutionary mode and tempo

Gould³⁵ suspected that “punctuated gradualism” was rare. In contrast, punctuated equilibrium (with speciation by “budding cladogenesis”³⁶) is thought to be more common, but demonstrating it requires the verified contemporaneity and persistence of both the ancestral and daughter species. As Gould³⁵ noted, “We can distinguish the punctuations of rapid anagenesis from those of branching speciation by invoking the eminently testable criterion of ancestral survival following the origin of a descendant species. If the ancestor survives, then the new species has arisen by branching. If the ancestor does not survive, then we must count the case either as indecisive, or as good evidence for rapid anagenesis—but, in any instance, not as evidence for punctuated equilibrium.” (p. 795).

These requirements have rarely been met among fossil hominids. For the origin of *Australopithecus*, phyletic evolution with a burst of rapid directional change during the 200,000-yr period between 4.4 and 4.2 Myr ago remains plausible given the geographic, temporal and morphological relationships of *Ar. ramidus* and *Au. anamensis* and our understanding of primate dental anatomy and development³⁷. Indeed, given the available evidence, the origin of *Australopithecus* could well turn out to be a case of “punctuated gradualism”³⁸ or “punctuated anagenesis”³⁹ rather than rectangular evolution *sensu* Stanley⁴⁰. Only the recovery of additional fossils from dated contexts will allow a more accurate and precise determination of the mode and tempo of early hominid evolution on the African continent.

Early hominid adaptation

Whatever the geometry of early hominid phylogeny, diagnostic

megadontia and related dentognathic morphology of *Australopithecus* herald its appearance at or before 4.2 Myr ago. Its masticatory apparatus appears better adapted to a more heavily chewed diet of tough and abrasive items^{41,42} than that of *Ardipithecus*. These phenotypic signals indicate an adaptive shift towards the exploitation of tougher and more abrasive food resources. This may signal an ‘ecological breakout’ involving niche expansion with intensified exploitation of more open African Pliocene habitats². Such habitats were evidently available even in the Late Miocene⁴³, but hominids older than *Australopithecus* apparently did not exploit them as intensively. Greater habitat and dietary specificity among the earliest hominids probably explains the difficulty of recovering them as fossils from pre-*Australopithecus* deposits, except in specific ecological circumstances.

The Asa Issie occurrence of *Au. anamensis* suggests that the habitat previously frequented by *Ardipithecus* continued to attract its more derived and probably more eurytopic descendant. Having crossed the threshold of megadontia, all *Australopithecus* subsequent to Asa Issie (the *Au. anamensis*–*Au. afarensis*–*Au. garhi* lineage; the *Au. aethiopicus*–*Au. boisei* lineage; *Au. africanus*; *Au. robustus*) continued to display hypertrophy of craniodental features, presumably evolved under natural selection involving intensified mastication. Species of the genus *Homo* violated this trend, but only subsequent to the appearance of stone tools.

The origin of *Australopithecus* between 4 and 5 Myr ago does not seem to correspond to any proxy signalling global climatic change^{44,45}. Neither is there an obvious pattern of an evolutionary pulse affecting other contemporary mammalian lineages. The triggers for the adaptive shift towards early hominid megadontia and the Pliocene origin of *Australopithecus* therefore remain elusive.

METHODS

Aramis vertebrate palaeontological locality 14 is approximately 9 m above the Kullunta basaltic tuff (KUBT) dated to 4.317 ± 0.055 Myr ago¹⁶ and about 50 m below another volcanic stratum, MA 94-55C, dated to 4.041 ± 0.060 Myr ago¹⁶. The hominid fossil maxilla ARA-VP-14/1 was found in 1994 in a palaeomagnetically reversed, orange–brown silty clay. The reversal immediately below the maxilla appears to be the base of chron C2Ar (ref. 46) at 4.21 Myr ago (recalculated; younger limit of chron is 3.61 Myr ago).

The ASI/HAN stratigraphic succession is dominated by fluvial deposits and minor interbedded bentonitic and glassy silicic and hydromagmatic basaltic tephra deposits. At the ASI-VP-2 site, the local stratigraphic sequence dips slightly to the west and is emplaced atop basalts and sediments locally representing the Gawto, Haradaso and Aramis members¹⁶.

All of the Asa Issie faunal collections and hominid remains described here are from a 2–3-m-thick zone of fluvial silty clays with root cast and pedogenic carbonates atop a widespread yellow–green altered hydromagmatic basaltic tephra (Fig. 1). Four samples spanning the 3 m of fossiliferous strata at the ASI-VP-2 locality are of reversed geomagnetic polarity, consistent with their placement in chron C2Ar. The upper portion of the underlying basaltic tephra includes a 15-cm-thick basaltic tuff that consists of fine- to medium-grained glass shards that are strongly calcite cemented and partially altered. It is a primary deposit and has homogeneous chemistry. We have dated it by ^{40}Ar – ^{39}Ar to 4.116 ± 0.074 Myr ago (MA02-13, weighted mean of the two plateau ages; integrated ages of 4.4 ± 0.4 Myr and 4.6 ± 0.2 Myr ago; Supplementary Table 4 and Supplementary Fig. 3), providing a maximum age for the fossils.

A widespread gastropod-bearing sandstone marker horizon overlies the hominid-bearing localities. In other portions of the Middle Awash study area a similar unit is capped by the Early Pliocene Cindery tuff dated by ^{40}Ar – ^{39}Ar to 3.88 ± 0.02 Myr ago^{16,47–49}. At Asa Issie this marker horizon is separated from younger deposits by a west-dipping, northwest–southeast-trending normal fault with ~30 m of displacement. This fault crosses ~100 m southwest of the ASI-VP-2 site, and the hanging wall contains a distinct 90–120-cm silicic tephra (MA00-19, MA00-20, MA00-21), the geochemical characteristics of which provide a firm correlation to the widespread 3.77-Myr VT-3/Wargolo tuff^{16,47–50}. From the cumulative evidence we conclude that the vertebrate fossils from the ASI-VP-2, ASI-VP-5 and HAN-VP-1 localities were embedded between 4.12 and 3.77 Myr ago, but much closer to the former based on stratigraphic proximity and biochronological considerations.

Received 22 September 2005; accepted 2 February 2006.

1. Asfaw, B. The Belohdelie frontal: New evidence of early hominid cranial morphology from the Afar of Ethiopia. *J. Hum. Evol.* **16**, 611–624 (1987).
2. White, T. D., Suwa, G. & Asfaw, B. *Australopithecus ramidus*, a new species of early hominid from Aramis, Ethiopia. *Nature* **371**, 306–312 (1994).
3. Haile-Selassie, Y. Late Miocene hominids from the Middle Awash, Ethiopia. *Nature* **412**, 178–181 (2001).
4. White, T. D. in *The Primate Fossil Record* (ed. Hartwig, W.) 407–417 (Cambridge Univ. Press, Cambridge, 2002).
5. Leakey, M. G., Feibel, C. S., McDougall, I. & Walker, A. New four-million-year-old hominid species from Kanapoi and Allia Bay, Kenya. *Nature* **376**, 565–571 (1995).
6. Leakey, M. G., Feibel, C. S., McDougall, I., Ward, C. & Walker, A. New specimens and confirmation of an early age for *Australopithecus anamensis*. *Nature* **393**, 62–66 (1998).
7. Ward, C., Leakey, M. & Walker, A. C. The new hominid species *Australopithecus anamensis*. *Evol. Anthropol.* **7**, 197–205 (1999).
8. Ward, C., Leakey, M. G. & Walker, A. Morphology of *Australopithecus anamensis* from Kanapoi and Allia Bay, Kenya. *J. Hum. Evol.* **41**, 255–368 (2001).
9. Haile-Selassie, Y., Suwa, G. & White, T. D. Late Miocene teeth from Middle Awash, Ethiopia, and early hominid dental evolution. *Science* **303**, 1503–1505 (2004).
10. Senut, B. *et al.* First hominid from the Miocene (Lukeino Formation, Kenya). *C.R. Acad. Sci. (Paris)* **332**, 134–144 (2000).
11. Pickford, M., Senut, B., Grommery, D. & Treil, J. Bipedalism in *Orrorin tugenensis* revealed by its femora. *C.R. Palevol* **1**, 191–203 (2002).
12. Brunet, M. *et al.* A new hominid from the upper Miocene of Chad, central Africa. *Nature* **418**, 145–151 (2002).
13. Ohman, J., Lovejoy, C. O. & White, T. D. Questions about *Orrorin* femur. *Science* **307**, 845b (2005).
14. White, T. D. Early hominid femora: The inside story. *Palevol* (in the press).
15. White, T. D. in *Paleoclimate and Evolution, with Emphasis on Human Origins* (eds Vrba, E., Denton, G., Partridge, T. & Burkle, L.) 369–384 (Yale Univ. Press, New Haven, 1995).
16. Renne, P. R., WoldeGabriel, G., Hart, W. K., Heiken, G. & White, T. D. Chronostratigraphy of the Miocene-Pliocene Sagantole Formation, Middle Awash Valley, Afar Rift, Ethiopia. *Geol. Soc. Am. Bull.* **111**, 869–885 (1999).
17. Renne, P. R. *et al.* Intercalibration of standards, absolute ages and uncertainties in $^{40}\text{Ar}/^{39}\text{Ar}$ dating. *Chem. Geol. Isot. Geosci. Sect.* **145**, 117–152 (1998).
18. Hlusko, L. J. A new large Pliocene colobine species (Mammalia: Primates) from Asa Issie, Ethiopia. *Geobios* **39**, 57–69 (2006).
19. Cerling, T. E., Bowman, J. R. & O'Neil, J. R. An isotopic study of a fluvial-lacustrine sequence: The Plio-Pleistocene Koobi Fora sequence, East Africa. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **63**, 335–356 (1988).
20. Levin, N. E., Quade, J., Simpson, S. W., Semaw, S. & Rogers, M. Isotopic evidence for Plio-Pleistocene environmental change at Gona, Ethiopia. *Earth Planet. Sci. Lett.* **219**, 93–110 (2004).
21. WoldeGabriel, G. *et al.* Geology and palaeontology of the Late Miocene Middle Awash Valley, Afar Rift, Ethiopia. *Nature* **412**, 175–178 (2001).
22. Haile-Selassie, Y. *et al.* Mio-Pliocene mammals from the Middle Awash, Ethiopia. *Geobios* **37**, 536–552 (2004).
23. Wynn, J. G. Paleosols, stable carbon isotopes, and paleoenvironmental interpretation of Kanapoi, Northern Kenya. *J. Hum. Evol.* **39**, 411–432 (2000).
24. Flessa, K. *et al.* The Geological Record of Ecological Dynamics: Understanding the Biotic Effects of Future Environmental Change (National Academy of Sciences, Washington, 2005).
25. Kono, R. T. Molar enamel thickness and distribution patterns in extant great apes and humans: New insights based on a 3-dimensional whole crown perspective. *Anthropol. Sci.* **112**, 121–146 (2004).
26. Suwa, G. & Kono, R. A micro-CT based study of linear enamel thickness in the mesial cusp section of human molars: Reevaluation of methodology and assessment of within-tooth, serial, and individual variation. *Anthropol. Sci.* **113**, 273–289 (2005).
27. Coffing, K., Feibel, C., Leakey, M. & Walker, A. Four-million-year-old hominids from east Lake Turkana, Kenya. *Am. J. Phys. Anthropol.* **93**, 55–65 (1994).
28. Schoeninger, J. J., Reeser, H. & Hallin, K. Paleoenvironment of *Australopithecus anamensis* at Allia Bay, East Turkana, Kenya: Evidence from mammalian herbivore enamel stable isotopes. *J. Anthropol. Archaeol.* **22**, 200–207 (2003).
29. Semaw, S. *et al.* Early Pliocene hominids from Gona, Ethiopia. *Nature* **433**, 301–305 (2005).
30. Wood, B. & Richmond, B. G. Human evolution: Taxonomy and paleobiology. *J. Anat. (Lond.)* **196**, 19–60 (2000).
31. Kappelman, J. *et al.* Age of *Australopithecus afarensis* from Fejej, Ethiopia. *J. Hum. Evol.* **30**, 139–146 (1996).
32. Haile-Selassie, Y. & Asfaw, B. A newly discovered early Pliocene hominid-bearing paleontological site in the Mulu Basin, Ethiopia. *Am. J. Phys. Anthropol.* **30** (suppl.), 170 (2000).
33. Wood, B. The oldest hominid yet. *Nature* **371**, 280–281 (1994).
34. Stern, J. T. & Susman, R. L. The locomotor anatomy of *Australopithecus afarensis*. *Am. J. Phys. Anthropol.* **60**, 279–317 (1983).
35. Gould, S. J. *The Structure of Evolutionary Theory* (Belknap, Harvard, Cambridge, 2002).
36. Wagner, P. The quality of the fossil record and the accuracy of phylogenetic inferences about sampling and diversity. *Syst. Biol.* **49**, 65–86 (2000).
37. Hlusko, L. J., Suwa, G., Kono, R. T. & Mahaney, M. C. Genetics and the evolution of primate enamel thickness: A baboon model. *Am. J. Phys. Anthropol.* **124**, 223–233 (2004).
38. Malmgren, B. A., Berggren, W. A. & Lohman, G. P. Evidence for punctuated gradualism in the Late Neogene *Globorotalia tumida* lineage of planktonic foraminifera. *Paleobiology* **9**, 377–389 (1983).
39. McKinney, M. L. & McNamara, K. J. *Heterochrony: The Evolution of Ontogeny* (Plenum, New York, 1991).
40. Stanley, S. M. *Macroevolution: Pattern and Process* (W. H. Freeman, San Francisco, 1979).
41. Teaford, M. F. & Ungar, P. S. Diet and the evolution of the earliest human ancestors. *Proc. Natl Acad. Sci. USA* **97**, 13506–13511 (2000).
42. Macho, G. A., Shimizu, D., Jiang, Y. & Spears, I. R. *Australopithecus anamensis*: A finite-element approach to studying the functional adaptations of extinct hominins. *Anat. Rec.* **283A**, 310–318 (2005).
43. Kingston, J. D., Marino, B. D. & Hill, A. Isotopic evidence for Neogene hominid paleoenvironments in the Kenya Rift Valley. *Science* **264**, 955–959 (1994).
44. DeMenocal, P. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth Planet. Sci. Lett.* **220**, 3–24 (2004).
45. Vrba, E. S. Mass turnover and heterochrony events in response to physical change. *Paleobiology* **31**, 157–174 (2005).
46. Cande, S. C. & Kent, D. V. Revised calibration of the geomagnetic polarity timescale for the Late Cretaceous and Cenozoic. *J. Geophys. Res.* **100**, 6093–6095 (1995).
47. Hall, C. M., Walter, R. C., Westgate, J. A. & York, D. Geochronology, stratigraphy and geochemistry of Cinderly Tuff in Pliocene hominid-bearing sediments of the Middle Awash, Ethiopia. *Nature* **308**, 26–31 (1984).
48. White, T. D. *et al.* New discoveries of *Australopithecus* at Maka, Ethiopia. *Nature* **366**, 261–265 (1993).
49. Hart, W. K., Walter, R. C. & WoldeGabriel, G. Tephra sources and correlations in Ethiopia: Application of elemental and neodymium isotope data. *Quatern. Int.* **13/14**, 77–86 (1992).
50. Haileab, B. & Brown, F. H. Turkana Basin–Middle Awash Valley correlations and the age of the Sagantole and Hadar Formations. *J. Hum. Evol.* **22**, 453–468 (1992).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the National Science Foundation (including the Revealing Hominid Origins Initiative/HOMINID program), the Institute of Geophysics and Planetary Physics of the University of California at Los Alamos National Laboratory (LANL), the Japan Society for the Promotion of Science, the Fondation Singer-Polignac, and the Philip and Elaine Hampton Fund for Faculty International Initiatives at Miami University for financial support of field and laboratory research. The Earth and Environmental Sciences Division Electron Microprobe laboratory at LANL assisted with access and use. We thank A. Ademassu, W. Amerga, A. Amzaye, A. Asfaw, G. Assefa, F. Bibi, M. Black, D. Brill, K. Brudvik, M. Chalachew, M. Chubachi, W. Demisse (in memoriam), N. Eldredge, H. Elema, E. Gülec, M. Haydera, R. Jabbour, A.-R. Jaouni, K. Kairento, F. Kaya, K. Kimeu, B. Kyongo, D. Kubo, W. Liu, S. Mahieu, W. Mangao, M. McCollum, E. Mekonnen, W. Mihel, L. Morgan, C. Pehlevan, P. Reno, G. Richards, B. Rosenman, M. Serrat, A. Shabel, L. Smeenk, B. Tegengne, A. Terrazas and S. Yoseph for fieldwork, laboratory work and discussion, and M. Leakey for collections access. We thank the Ministry of Youth, Sports and Culture, the Authority for Research and Conservation of the Cultural Heritage, and the National Museum of Ethiopia for permissions and facilitation. We also thank the Afar Regional Government, the Afar people of the Middle Awash, and many others for contributing directly to the research efforts.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.W. (timwhite@berkeley.edu).

ARTICLES

Chemokines enhance immunity by guiding naive CD8⁺ T cells to sites of CD4⁺ T cell-dendritic cell interaction

Flora Castellino^{1*†}, Alex Y. Huang^{1*}, Grégoire Altan-Bonnet^{1†}, Sabine Stoll^{1†}, Clemens Scheinecker^{1†} & Ronald N. Germain¹

CD8⁺ T cells have a crucial role in resistance to pathogens and can kill malignant cells; however, some critical functions of these lymphocytes depend on helper activity provided by a distinct population of CD4⁺ T cells. Cooperation between these lymphocyte subsets involves recognition of antigens co-presented by the same dendritic cell, but the frequencies of such antigen-bearing cells early in an infection and of the relevant naive T cells are both low. This suggests that an active mechanism facilitates the necessary cell-cell associations. Here we demonstrate that after immunization but before antigen recognition, naive CD8⁺ T cells in immunogen-draining lymph nodes upregulate the chemokine receptor CCR5, permitting these cells to be attracted to sites of antigen-specific dendritic cell-CD4⁺ T cell interaction where the cognate chemokines CCL3 and CCL4 (also known as MIP-1 α and MIP-1 β) are produced. Interference with this actively guided recruitment markedly reduces the ability of CD4⁺ T cells to promote memory CD8⁺ T-cell generation, indicating that an orchestrated series of differentiation events drives nonrandom cell-cell interactions within lymph nodes, optimizing CD8⁺ T-cell immune responses involving the few antigen-specific precursors present in the naive repertoire.

CD4⁺ and CD8⁺ T cells represent two major lymphocyte subsets involved in adaptive immune responses. Recently, there has been renewed interest in understanding how CD4⁺ T cells contribute to the immune activities of CD8⁺ T cells and how such cooperation is mediated^{1,2}. There is growing agreement that CD4⁺ T cells have an especially important role in assisting the development of long-term CD8⁺ T-cell memory³ and that the antigenic ligands for the CD4⁺ and CD8⁺ T cells involved in such help-dependent responses must be co-presented by the same antigen-presenting dendritic cell². However, two temporally distinct models have been offered for the relevant cell interactions. One suggests that both lymphocyte types bind simultaneously to a single dendritic cell, producing a ternary cluster within which local cytokine production (for example, interleukin (IL)-2) or direct interaction via CD40-CD40L are the key elements of intercellular cooperation^{4,5}. Because CD4⁺ and CD8⁺ T cells specific for a given antigen are infrequent in the naive lymphocyte pool, others have argued that it is very unlikely for two rare T cells to interact simultaneously with one of the small number of relevant dendritic cells present early in an infection⁶. This view is supported by recent imaging studies concluding that CD4⁺ and CD8⁺ T cells interact in a random fashion with antigen-bearing dendritic cells within lymph nodes⁷⁻¹¹. Such considerations led to a 'sequential model' postulating that the interaction of a CD4⁺ T cell with an antigen-presenting dendritic cell induces differentiation of the latter ('licensing') that supports CD8⁺ T-cell priming even after the CD4⁺ T cell has dissociated¹²⁻¹⁵.

This sequential model does not consider the possibility that naive

CD8⁺ T cells could be actively attracted to those few dendritic cells in a lymph node presenting antigen to CD4⁺ T cells. In addition, recent evidence indicates that antigen-dependent CD4⁺ T-cell interactions with dendritic cells in lymph nodes can last for several hours¹⁶, providing a substantial window of time during which a third cell, even if rare, might find a pre-existing CD4⁺ T cell-dendritic cell pair. Furthermore, suitably activated dendritic cells produce chemokines that could attract CD8⁺ T cells expressing appropriate receptors¹⁷⁻¹⁹. On the basis of these considerations, we undertook a series of experiments to examine whether cell-cell interactions in lymph nodes, in an immunogenic environment, were influenced by chemokines, and whether nonrandom associations of CD8⁺ T cells with antigen-activated dendritic cells and/or dendritic cell-CD4⁺ T-cell clusters occurred during immune responses.

Activated CD4⁺ T cells cause CD8⁺ T-cell accumulation

If naive CD8⁺ T cells were attracted to sites of antigen-dependent CD4⁺ T cell-dendritic cell interaction, this might result in their accumulation in lymph nodes where such CD4⁺ T-cell activation was occurring. To examine this possibility, OT-I (CD8⁺)²⁰ and OT-II (CD4⁺)²¹ T-cell receptor (TCR) transgenic T cells of known antigen specificity were co-transferred into immunocompetent recipients. The mice were then immunized subcutaneously on two opposite sides and the distribution of OT-I T cells in the two draining lymph nodes compared 40 h later. When mice were immunized with the same vaccine formulation on both sides, OT-I T cells distributed equally to the two draining lymph nodes (Fig. 1). If the peptide

¹Lymphocyte Biology Section, Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

[†]Present addresses: Chiron Vaccines, Via Fiorentina 1, Siena 53100, Italy (F.C.); cBio, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA (G.A.-B.); Department of Dermatology, University of Mainz, Mainz 55131, Germany (S.S.); Department of Rheumatology, Internal Medicine III, Medical University of Vienna, General Hospital of Vienna, Währinger Gürtel 18-20, A-1090 Vienna, Austria (C.S.).

*These authors contributed equally to this work.

antigen for the OT-II CD4⁺ T cells, OVA 323–339, was given only on one side, however, there was a highly significant 1.5–2-fold greater accumulation of OT-I CD8⁺ T cells in the lymph node on that side as compared to the contralateral node (Fig. 1). The accumulation of OT-I T cells was independent of the presence of their cognate antigen OVA 257–264 and of major histocompatibility complex (MHC) class I molecule expression by the dendritic cells (Fig. 1). Even when OT-II cells were not co-transferred, immunization with OVA 323–339 resulted in a modest (1.3-fold) increase in OT-I accumulation, presumably due to the activity of endogenous OVA 323–339-specific CD4⁺ T cells. Similarly, in the presence of OT-II cells, lipopolysaccharide (LPS)-activated, OVA 323–339-pulsed bone-marrow-derived dendritic cells promoted a greater accumulation of OT-I T cells in the draining lymph node than did LPS-treated dendritic cells lacking OVA 323–339 (Fig. 1). These findings are unlikely to result solely from ‘lymph node shut-down’, the transient, proportional accumulation of all lymphocyte subsets in a lymph node draining a site of inflammation^{22–24}. CD8⁺ T cells showed an approximately 2-fold greater accumulation than endogenous CD4⁺ T cells on the side where OVA 323–339-stimulated OT-II T-cell activation was occurring (data not shown). Together, these findings indicate that naive CD8⁺ T cells accumulate in lymph nodes where CD4⁺ T cells are undergoing antigen-specific activation, without requiring TCR-dependent activation of the CD8⁺ T cells.

Nonrandom CD8⁺ T cell-dendritic cell interactions

To determine whether this accumulation was accompanied by enhanced physical interactions between naive CD8⁺ T cells and dendritic cells presenting antigen to CD4⁺ T cells as compared to dendritic cells lacking such antigen, we performed intravital two-photon microscopy to track individual T cells and dendritic cells within the lymphoid environment. LPS-activated dendritic cells were either left unpulsed or pulsed with OVA 323–339, then labelled with distinct fluorescent dyes and injected subcutaneously. Six to twelve hours later naive CD4⁺ and CD8⁺ T cells, stained with distinct dyes, were introduced intravenously. Sixteen to thirty hours after T-cell

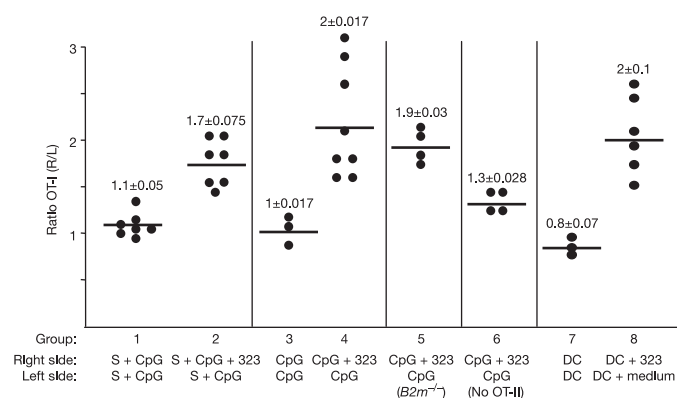


Figure 1 | CD8⁺ T cells accumulate in lymph nodes in which antigen-dependent CD4⁺ T cell-dendritic cell interactions are occurring. Unless otherwise specified (No OT-II), mice were co-injected intravenously with OT-I and OT-II T cells, then immunized subcutaneously with either OVA 323–339 (323), OVA 257–264 (S) and/or CpGs in alum (groups 1–6). Alternatively, LPS-activated bone-marrow-derived dendritic cells were pulsed with OVA 323–339 (DC + 323) or left unpulsed (DC + medium), then injected subcutaneously followed by OT-I and OT-II T cells intravenously (groups 7–8). OT-I distributions in draining lymph nodes were determined 40 h after immunization. Data represent the relative accumulation of CD8⁺ T cells in the right versus the left lymph nodes (R/L). Each dot is the R/L ratio from an individual mouse, with the average ± standard error shown above. Groups 1 versus 2, $P = 0.00058$; 3 versus 4, $P = 0.012$; 3 versus 5, $P = 0.00095$; 3 versus 6, $P = 0.021$; 7 versus 8, $P = 0.024$.

transfer, the draining lymph nodes were imaged in living animals. Because the measured frequency of labelled CD8⁺ T-cell contacts with labelled dendritic cells will vary with imaging volume, time and cell densities, we normalized the primary imaging data (Fig. 2a, b; see also Supplementary Movies 1 and 2) to generate a parameter termed the hit rate constant or K_{on} . This derived parameter represents the normalized frequency of CD8⁺ T-cell contacts per dendritic cell in an imaging data set (see Supplementary Information for details). We first derived the hit rate constants in each experiment for each T cell-dendritic cell pair studied and used these data for cross-experiment comparisons (Fig. 2c; see also Supplementary Table 1). Pairs of hit rate constants derived from a single imaging data set were also used to produce a hit rate ratio, which represents the relative rates of CD8⁺ T-cell contact with OVA 323–339-bearing versus OVA 323–339-free dendritic cell populations co-localized within an individual lymph node (Fig. 2d). Analysis based on simultaneous imaging of all relevant cell populations in a single lymph node can be critical to proper data interpretation^{25,26}. Using these methods, we found that when OT-II T cells were present, the hit rate constant for OT-I T cells contacting OVA 323–339-pulsed dendritic cells was 3.5-fold higher than that involving unpulsed dendritic cells (Fig. 2d; see also Supplementary Table 1). Similar results were obtained when polyclonal CD8⁺ T cells were used instead of OT-I cells (Fig. 2d; see also Supplementary Table 1). In agreement with the accumulation data reported in Fig. 1, a smaller but significant increase (2.3-fold) in the hit rate constant with OVA 323–339-pulsed dendritic cells as compared to unpulsed dendritic cells was observed even when OT-II T cells were not transferred (Fig. 2d; see also Supplementary Table 1).

CCL3/CCL4-CCR5 binding promotes CD8⁺ T-cell accumulation

One possible explanation for the preceding results is chemokine-dependent attraction^{17,18} of CD8⁺ T cells to sites of CD4⁺ T-cell activation. We therefore attempted to interfere with CD4⁺ T-cell-dependent lymph node accumulation of naive CD8⁺ T cells by injecting blocking antibodies against candidate chemokines. Antibodies against CCL3 or CCL4 consistently reduced the accumulation of naive OT-I T cells in the lymph nodes with activated OT-II T cells (Fig. 3a). Other tested chemokine-blocking antibodies had minimal or no effect on lymph node recruitment of naive CD8⁺ T cells (Fig. 3a).

Consistent with the finding that blocking CCL3 and CCL4 reduced CD4⁺ T-cell-dependent CD8⁺ T-cell accumulation, lymph nodes draining vaccine sites containing OVA 323–339 produced more CCL3 and CCL4 than those exposed to adjuvant alone (Supplementary Fig. 1a, b). Both dendritic cells and CD4⁺ T cells produced CCL3 and CCL4 upon activation *in vitro* or *in vivo* (Supplementary Fig. 1c–e; see also ref. 27), implying that CD4⁺ T cell-dendritic cell clusters would be most attractive to CD8⁺ T cells. In agreement with previous studies^{28,29}, supernatant from vaccine-draining lymph nodes showed evidence of CCL3 and CCL4 association, offering a potential explanation for the non-redundant ability of antibodies against each chemokine to interfere with CD8⁺ T-cell accumulation (Fig. 3a; see also Supplementary Fig. 1f).

The contribution of CCL3 and CCL4 to CD8⁺ T-cell accumulation was surprising, given that naive CD8⁺ T cells from unmanipulated mice do not express the relevant chemokine receptor (CCR5) and do not migrate in response to CCL4 *in vitro* (Supplementary Fig. 2b). CCR5 was detectable, however, on a fraction of the OT-I T cells recovered from lymph nodes draining sites of immunization, with the percentage of positive cells varying with the stimulus (Supplementary Fig. 3a, b). OT-I T cells recovered from non-draining lymph nodes did not express CCR5 (data not shown). OT-I T-cell expression of CCR5 was not dependent on TCR signalling, in contrast to OT-II T cells (Supplementary Fig. 3c), although such expression could be further increased by TCR stimulation (Supplementary Fig. 3a). Consistent with these staining data, OT-I cells purified from draining lymph nodes 40 h after vaccination migrated

in vitro in response to CCL4 (Supplementary Fig. 2c). Finally, wild type but not *Ccr5*^{-/-} CD8⁺ T cells accumulated efficiently in lymph nodes containing activated CD4⁺ T cells (Fig. 3b), confirming the critical role of CCL3/CCL4–CCR5 interactions in this process.

Chemokines enhance CD8⁺ T-cell contacts with dendritic cells

We next studied wild-type and *Ccr5*^{-/-} polyclonal CD8⁺ T cells using dynamic intravital imaging, to determine whether the preferential interaction of CD8⁺ T cells with dendritic cells presenting antigen to CD4⁺ T cells could be explained by a chemokine influence (Supplementary Movie 4). In the presence of OT-II T cells, wild-type CD8⁺ T cells had a hit rate ratio of 3.97 when comparing interactions with OVA 323–339-pulsed dendritic cells to those involving unpulsed dendritic cells. In contrast, *Ccr5*^{-/-} CD8⁺ T cells in the same lymph nodes had a hit rate ratio of 1.14 for their interactions with the same two dendritic cell populations (Fig. 3c; see also Supplementary Table 1). Antibody neutralization of CCL3 and CCL4 *in vivo* eliminated the difference in the hit rate constants of wild-type CD8⁺ T cells interacting with OVA 323–339-pulsed versus unpulsed dendritic cells in the presence of OT-II T cells (Fig. 3d; see also Supplementary Table 1).

To determine the contributions of chemokinesis and chemotaxis to these chemokine-dependent effects on CD8⁺ T-cell contact with dendritic cells, we analysed the migratory behaviour of wild-type and *Ccr5*^{-/-} CD8⁺ T cells in the vicinity of OVA 323–339-pulsed dendritic cells. The dendrites of dendritic cells physically scan a region approximately 50 μm in diameter⁹. Therefore, to analyse

T-cell movement in a manner relevant to the hit rate data presented above, we determined the frequency with which T cells starting more than 30 μm away from a sphere of 25- μm radius centred on an isolated, LPS-activated, OVA 323–339-pulsed dendritic cell reached the border of this sphere (Supplementary Fig. 5). The requirement of tracked T cells beginning their approach to the dendritic cell at a substantial distance from the potential point of contact was necessary to allow an adequate analysis of speed, persistence and directionality. Notably, only 6 out of 518 (1.1%) *Ccr5*^{-/-} T cells as compared with 24 out of 553 (4.7%) wild-type T cells reached the border of an OVA 323–339-pulsed dendritic cell from the neighbouring region of the lymph node during a typical 30-min imaging session (Fig. 4a, b). The relative frequency (4.3) with which wild-type versus *Ccr5*^{-/-} T cells moved long distances towards a dendritic cell closely approximated the hit rate ratio reported in Fig. 3c for the two lymphocyte types. To look for a chemokinetic effect, the instantaneous speed of the T cells that came within 25 μm of the dendritic cell centre was determined (Fig. 4c). The wild-type T-cell subset had a 2.5-fold higher frequency of individual steps showing a fast speed (migration at $\geq 10 \mu\text{m min}^{-1}$) as compared with the *Ccr5*^{-/-} subset (25% versus 10%, respectively). The rapid steps of wild-type T cells showed a directional bias towards the OVA 323–339-pulsed dendritic cell, consistent with chemotactic behaviour (Fig. 4d; see also Supplementary Fig. 6). None of these differences was seen in lymph nodes that were not draining a vaccine site (Supplementary Fig. 7). These data indicate that a fraction of CD8⁺ T cells show chemokine-guided migration and acceleration towards activated dendritic cells that are

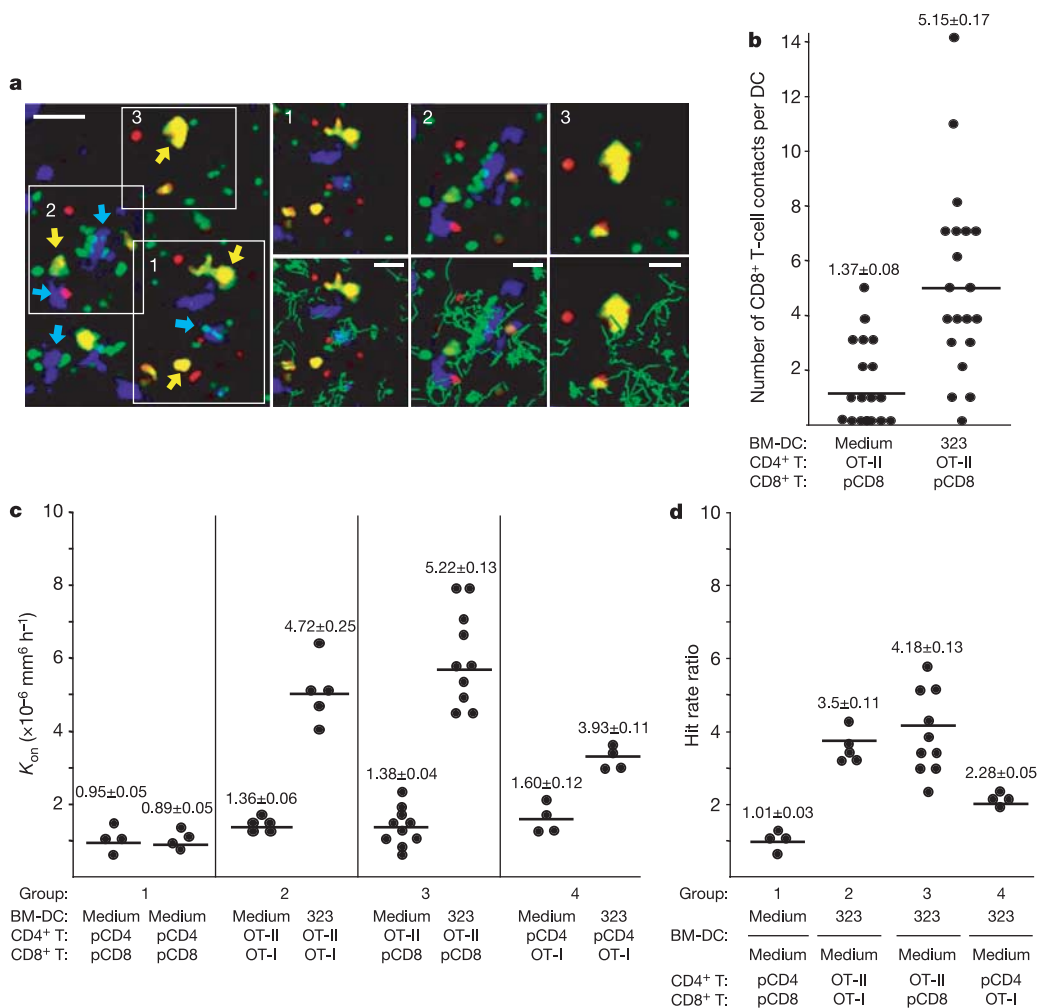


Figure 2 | Increased frequency of cellular contacts between CD8⁺ T cells and dendritic cells capable of antigen-dependent interaction with CD4⁺ T cells. **a**, Left: OT-II CD4⁺ T cells (red) and polyclonal CD8⁺ T cells (green) in the interfollicular region of a draining lymph node containing OVA 323–339-pulsed dendritic cells (blue/blue arrows) and medium-pulsed dendritic cells (yellow/yellow arrows). The top row of the right panel shows enlarged views. The bottom row of the right panel shows tracks (green lines) made by CD8⁺ T cells. Scale bar: left, 30 μm ; right, 15 μm . **b**, Absolute number of CD8⁺ T-cell contacts per dendritic cell (DC). Nineteen unpulsed dendritic cells and 20 OVA 323–339-pulsed dendritic cells were examined ($P < 0.0001$ for the comparison of the two contact frequencies). BM-DC, bone-marrow-derived dendritic cell. **c**, Hit rate constants (K_{on}) for each CD8⁺ T cell–dendritic cell pair. Intragroup comparisons: 1, $P = 0.89$; 2, $P = 0.0079$; 3, $P < 0.0001$; 4, $P = 0.029$. Intergroup comparisons: 1 versus 2, $P = 0.11$; 1 versus 3, $P = 0.24$; 1 versus 4, $P = 0.029$. **d**, Hit rate ratios. All values are expressed as average $\pm$ standard error. pCD4, polyclonal CD4⁺ T cells; pCD8, polyclonal CD8⁺ T cells. Group 1 versus 2, $P < 0.0001$; 1 versus 3, $P = 0.002$; 1 versus 4, $P = 0.0049$; 2 versus 3, $P = 0.44$; 2 versus 4, $P = 0.016$; 3 versus 4, $P = 0.002$.

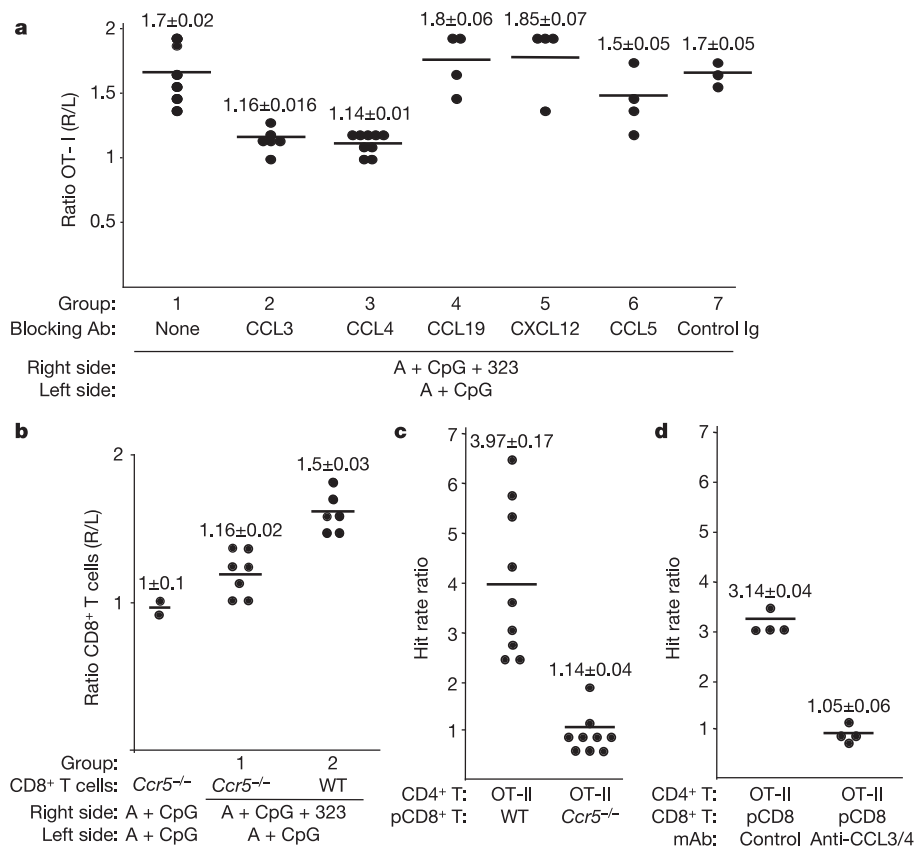


Figure 3 | Role of CCL3 and CCL4 in the CD4⁺ T cell-dendritic cell-dependent accumulation of CD8⁺ T cells within lymph nodes draining an immunization site. **a**, CCL3 and CCL4 neutralization interferes with OT-I accumulation in lymph nodes containing antigen-activated CD4⁺ T cells. Results are expressed as in Fig. 1. Group 1 versus 2, $P = 0.002$; 1 versus 3, $P = 0.0004$; 1 versus 4, $P = 0.48$; 1 versus 5, $P = 0.20$; 1 versus 6, $P = 0.26$; 1 versus 7, $P = 0.90$. **b**, *Ccr5*^{-/-} CD8⁺ T cells fail to accumulate in lymph nodes containing antigen-activated CD4⁺ T cells. Group 1 versus 2, $P = 0.0047$. **c**, *Ccr5*^{-/-} CD8⁺ T cells fail to show preferential contact with dendritic cells presenting antigen to CD4⁺ T cells. Hit rate constants for each CD8⁺ T cell-dendritic cell pair were derived from data sets involving imaging of the draining lymph node 16–30 h after T-cell injection (Supplementary Table 1). The graph represents nine independent imaging data sets from four different experiments ($P < 0.0001$ for the comparison between wild-type and *Ccr5*^{-/-} cells). **d**, CCL3 and CCL4 neutralization interferes with preferential contact of OT-I cells with dendritic cells presenting antigen to CD4⁺ T cells. The hit rate ratios were determined as described in **c** from four independent data sets. $P = 0.029$ for the comparison between control and antibody-blocked groups. A, alum.

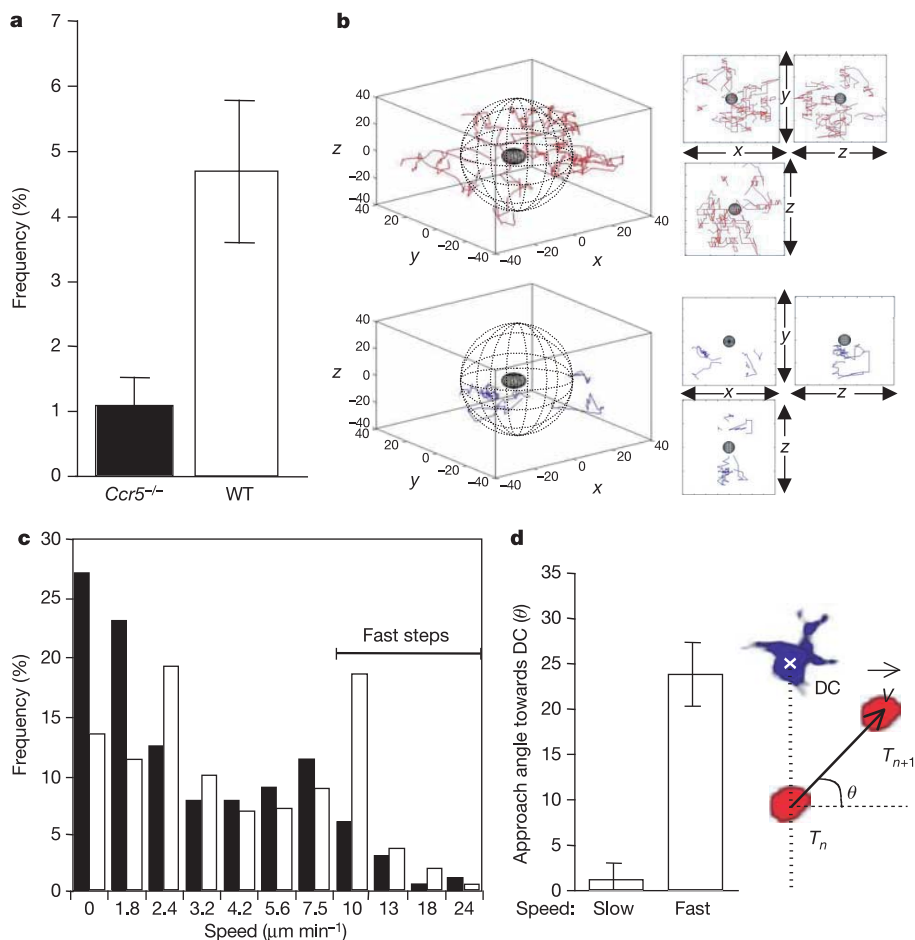


Figure 4 | Wild-type CD8⁺ T cells exhibit chemokine-directed rapid jumps and directional migration towards OVA 323-339-pulsed dendritic cells. **a**, *Ccr5*^{-/-} T cells show a reduced capacity as compared to wild-type T cells to reach a nearby dendritic cell presenting antigen to CD4⁺ T cells ($P = 0.009$). Data are from Supplementary Fig. 4. **b**, Trajectories of wild-type (red tracks) and *Ccr5*^{-/-} (blue tracks) T cells showing a wider area of distribution for wild-type tracks compared to *Ccr5*^{-/-} tracks around the centroid of the dendritic cell (centre globe). The dotted sphere (25-μm radius) represents the sweep volume of dendritic cell dendrites. Units are in micrometres. **c**, Distribution of instantaneous T-cell speed, v , in the vicinity of activated dendritic cells. *Ccr5*^{-/-} T-cell steps are indicated by black bars ($n = 171$); wild-type T-cell steps are indicated by white bars ($n = 572$). **d**, T-cell approach angles, θ , towards dendritic cells are $1.3 \pm 1.8^\circ$ for slow steps ($n = 441$) and $23.7 \pm 3.5^\circ$ for fast steps ($n = 131$) among wild-type cells ($P = 0.006$).

currently or have previously engaged in antigen-specific interactions with CD4⁺ T cells. Our findings are consistent with a recent imaging analysis of positively selecting thymocytes that also revealed a relationship between speed and directionality³⁰. The chemokine-mediated effect we describe here can account for the formation of ternary clusters of CD4⁺ T cells, CD8⁺ T cells and dendritic cells observed at both high precursor frequency (Supplementary Fig. 4a and Supplementary Movie 5) and with precursor numbers closer to the physiological range (Supplementary Fig. 4b), in agreement with a previous report of ternary cell clusters in immunohistochemical studies of disrupted lymph node cell preparations³¹.

Chemokine blockade interferes with CD8⁺ T-cell memory

To investigate the functional significance of this chemokine-mediated clustering, we examined the effect of blocking CCL3 and CCL4 action on acutely activated versus memory CD8⁺ T-cell number and effector capacity. Mice were adoptively transferred with OT-I and OT-II T cells and immunized in the presence or absence of OVA 323–339 and the presence or absence of neutralizing antibodies to CCL3 and CCL4 or isotype-matched control antibodies. Interfering with CCL3 and CCL4 function at the time of priming had a minimal effect on either the acute expansion or the acquisition of effector functions of OT-I T cells (Fig. 5). However, this manipulation eliminated the enhancing effect of CD4⁺ T cells on both the number and the effector activity of memory CD8⁺ OT-I T cells present several weeks later (Fig. 5). The lack of an effect of chemokine blockade on acute responses or on the generation of CD4⁺ T-cell-independent memory argues against CCL3 and CCL4 neutralization simply affecting early delivery of co-stimulatory

signals involved in CD8⁺ T-cell activation³² or late trafficking of activated cells to sites where survival factors are available. This conclusion is further supported by data showing that the contribution of CD4⁺ T-cell help to the generation of committed pre-memory CD8⁺ T cells within 7 days of immunization is eliminated by blocking CCL3 and CCL4 (F. C., manuscript submitted). Together, these findings are most consistent with other evidence that CD4⁺ T-cell help acts early after immunization to enhance the pool of functional memory CD8⁺ T cells (programming)^{5,13,33,34} (F. C., manuscript submitted) and link the optimal delivery of CD4⁺ T-cell help *in vivo* to the action of CCL3 and CCL4 in fostering productive cell–cell interactions within the lymph node environment. We cannot, of course, formally exclude other roles for CCL3 and CCL4 in the development of a robust memory CD8⁺ response, in addition to the early effects on cell–cell interactions we describe here.

Conclusion

Adaptive immune responses depend on the activation and function of rare antigen-specific lymphocytes bearing clonally distributed receptors. Despite evidence that the immune system has evolved mechanisms for efficient delivery of antigen-bearing cells to sites of naive lymphocyte concentration³⁵, many recent models of the early events during immune responses are based on the concept of random cell–cell interactions once the antigen-laden cells have arrived in lymphoid tissues^{8,9,36}. Our findings indicate that inflammation leads to CCR5 expression by antigen-naïve CD8⁺ T cells, permitting them to be attracted actively to sites of productive CD4⁺ T cell–dendritic cell interaction. By enhancing the scanning of these well-stimulated dendritic cells, the immune system increases the chances that rare antigen-specific CD8⁺ T cells will be optimally activated even with limiting numbers of antigen-loaded dendritic cells and specific CD4⁺ T cells. Although the measured effect of chemokine-directed recruitment was only 3–5-fold at the initial stage of the immune response, blocking CCL3 and CCL4 function early after immunization eliminated most of the contribution of CD4⁺ T-cell help to long-term CD8⁺ T-cell memory. Early, nonrandom cell clustering promoted by the cascade of chemokine-mediated events described here thus provides a new perspective on the environment within which CD4⁺ T cells facilitate physiologically useful cell-mediated immune responses.

METHODS

See Supplementary Information for additional details on all methods and procedures listed here, as well as for details on chemokine secretion assays and intracellular staining, hit rate analysis, and chemotaxis/chemokinesis analysis.

Animals. Transgenic mice on the C57BL/6 *Rag2*^{−/−} background expressing the OT-I TCR³⁰, transgenic mice on a C57BL/6 background expressing the OT-II TCR³¹, and congenic Ly5.2 C57BL/6 mice were obtained from Taconic Farms. Wild-type C57BL/6, *Ccr5*^{−/−} and β 2-microglobulin-deficient mice (*B2m*^{−/−}) (C57BL/6 background) were purchased from the Jackson Laboratory. All animals were housed and handled according to NIH institutional guidelines under an approved animal protocol.

Cell preparation, adoptive transfer and immunization. T-cell populations were isolated from lymph nodes and spleens and enriched by negative selection. The resultant cells were adoptively transferred into recipient mice that were immunized by subcutaneous injection with alum admixed with either OVA 323–339 (10 μ g), OVA 257–264 (1 μ g) and/or CpGs (GCTAGACGTTAGGT and TCAACGTTGA; 20 μ g total)³⁷. Bone-marrow-derived dendritic cells were prepared in murine granulocyte–macrophage colony-stimulating factor (mGM-CSF) and mIL-4, mock unpulsed (medium) or pulsed with 10 μ g ml^{−1} OVA 323–339 in the presence of 20 ng ml^{−1} LPS, then stained and injected subcutaneously into the dorsum of the foot. Various dye-labelled T-cell populations were transferred 6–18 h later. For analysis involving *in vivo* chemokine neutralization, various chemokine-blocking antibodies or isotype-matched controls were injected intravenously at the time of immunization.

Flow cytometry. T cells (3×10^6 cells per population) were injected intravenously into Ly5.2 congenic mice. Twenty-four hours later recipient mice were

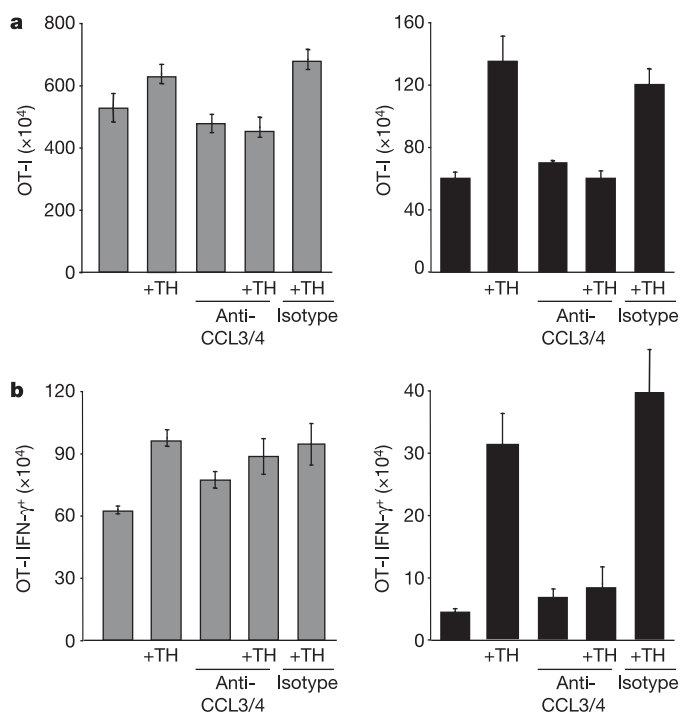


Figure 5 | Role of CCL3 and CCL4 in CD4⁺ T help for CD8⁺ T-cell memory. Mice adoptively transferred with naive OT-I and OT-II T cells were immunized with alum plus OVA 257–264 plus CpGs in the absence or presence of OVA 323–339 and neutralizing antibodies against CCL3 and CCL4. **a**, Numbers of OT-I T cells in lymph nodes and peripheral tissues 7 days (grey) and 21 days (black) later. **b**, Numbers of interferon (IFN)- γ secreting OT-I T cells after 4 h of *in vitro* antigen re-stimulation after recovery on day 7 (grey) and day 21 (black) after priming. Average $\pm$ standard error was calculated for data from three independent experiments. +TH, presence of antigen-activated CD4⁺ T cells; Isotype, isotype-matched control antibodies.

immunized subcutaneously. Forty hours after immunization CD8⁺ T cells in the draining popliteal lymph nodes were analysed by flow cytometry. Data are expressed as the ratio of CD8⁺ T-cell content (%) in the lymph node draining a vaccine formulation containing OVA 323–339 to the CD8⁺ T-cell content (%) in the contralateral lymph node draining a vaccine formulation without OVA 323–339.

Intravital two-photon imaging. Six to thirty hours after fluorescently labelled T-cell transfer, recipient mice were anaesthetized and two-photon intravital microscopy performed on the right popliteal lymph node using a modified surgical technique³⁸. Images were collected with typical voxel size of 0.91 × 0.91 × 3–5 μm and a volume dimension of 467 × 467 × 50–100 μm, unless otherwise noted. This volume collection was repeated every 15–40 s for up to 45 min to create a four-dimensional data set, processed using Imaris software (Bitplane) to create sequential two-dimensional maximum intensity projections and movies generated from these processed images.

Effect of CCL3/CCL4 blocking on CD8⁺ T-cell functional responses. OT-I and OT-II cells were co-transferred intravenously into congenic recipients. One day later the mice were immunized subcutaneously with alum admixed with OVA 257–264 (1 μg) and CpGs (20 μg) in the presence or absence of OVA 323–339 (10 μg) and in the presence or absence of neutralizing antibodies against CCL3 and CCL4 or isotype-matched control antibodies (50 μg in PBS per mouse). The mice were killed at different times after immunization, followed by cell isolation and staining for flow cytometry.

Statistics. All data were expressed as average values of experimental data points, and standard errors were determined using the calculated standard deviation of a data set divided by the number of data points within the data set. *P*-values between two non-normally distributed data sets were computed using the two-tailed Mann–Whitney *U*-test.

Received 22 November 2005; accepted 15 February 2006.

1. Germain, R. N. & Castellino, F. Co-operation between CD4⁺ and CD8⁺ T cells: when, where, and how. *Annu. Rev. Immunol.* (in the press).
2. Bevan, M. J. Helping the CD8⁺ T-cell response. *Nature Rev. Immunol.* **4**, 595–602 (2004).
3. Kaech, S. M., Wherry, E. J. & Ahmed, R. Effector and memory T-cell differentiation: implications for vaccine development. *Nature Rev. Immunol.* **2**, 251–262 (2002).
4. Keene, J. A. & Forman, J. Helper activity is required for the *in vivo* generation of cytotoxic T lymphocytes. *J. Exp. Med.* **155**, 768–782 (1982).
5. Bourgeois, C., Rocha, B. & Tanchot, C. A role for CD40 expression on CD8⁺ T cells in the generation of CD8⁺ T cell memory. *Science* **297**, 2060–2063 (2002).
6. Guerder, S. & Matzinger, P. A fail-safe mechanism for maintaining self-tolerance. *J. Exp. Med.* **176**, 553–564 (1992).
7. Miller, M. J., Wei, S. H., Parker, I. & Cahalan, M. D. Two-photon imaging of lymphocyte motility and antigen response in intact lymph node. *Science* **296**, 1869–1873 (2002).
8. Miller, M. J., Wei, S. H., Cahalan, M. D. & Parker, I. Autonomous T cell trafficking examined *in vivo* with intravital two-photon microscopy. *Proc. Natl Acad. Sci. USA* **100**, 2604–2609 (2003).
9. Miller, M. J., Hejazi, A. S., Wei, S. H., Cahalan, M. D. & Parker, I. T cell repertoire scanning is promoted by dynamic dendritic cell behavior and random T cell motility in the lymph node. *Proc. Natl Acad. Sci. USA* **101**, 998–1003 (2004).
10. Bousso, P. & Robey, E. Dynamics of CD8⁺ T cell priming by dendritic cells in intact lymph nodes. *Nature Immunol.* **4**, 579–585 (2003).
11. Mempel, T. R., Henrickson, S. E. & Von Andrian, U. H. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature* **427**, 154–159 (2004).
12. Bennett, S. R. *et al.* Help for cytotoxic-T-cell responses is mediated by CD40 signalling. *Nature* **393**, 478–480 (1998).
13. Sun, J. C. & Bevan, M. J. Defective CD8 T cell memory following acute infection without CD4 T cell help. *Science* **300**, 339–342 (2003).
14. Ridge, J. P., Di Rosa, F. & Matzinger, P. A conditioned dendritic cell can be a temporal bridge between a CD4⁺ T-helper and a T-killer cell. *Nature* **393**, 474–478 (1998).
15. Schoenberger, S. P., Toes, R. E., van der Voort, E. I., Offringa, R. & Melief, C. J. T-cell help for cytotoxic T lymphocytes is mediated by CD40–CD40L interactions. *Nature* **393**, 480–483 (1998).
16. Stoll, S., Delon, J., Brotz, T. M. & Germain, R. N. Dynamic imaging of T cell-dendritic cell interactions in lymph nodes. *Science* **296**, 1873–1876 (2002).
17. Sallusto, F. & Lanzavecchia, A. Understanding dendritic cell and T-lymphocyte traffic through the analysis of chemokine receptor expression. *Immunol. Rev.* **177**, 134–140 (2000).
18. Cyster, J. G. Lymphoid organ development and cell migration. *Immunol. Rev.* **195**, 5–14 (2003).
19. Proietto, A. I. *et al.* Differential production of inflammatory chemokines by murine dendritic cell subsets. *Immunobiology* **209**, 163–172 (2004).
20. Clarke, S. R. *et al.* Characterization of the ovalbumin-specific TCR transgenic line OT-I: MHC elements for positive and negative selection. *Immunol. Cell Biol.* **78**, 110–117 (2000).
21. Barnden, M. J., Allison, J., Heath, W. R. & Carbone, F. R. Defective TCR expression in transgenic mice constructed using cDNA-based α- and β-chain genes under the control of heterologous regulatory elements. *Immunol. Cell Biol.* **76**, 34–40 (1998).
22. MartIn-Fontecha, A. *et al.* Regulation of dendritic cell migration to the draining lymph node: impact on T lymphocyte traffic and priming. *J. Exp. Med.* **198**, 615–621 (2003).
23. Cahill, R. N., Frost, H. & Trnka, Z. The effects of antigen on the migration of recirculating lymphocytes through single lymph nodes. *J. Exp. Med.* **143**, 870–888 (1976).
24. Hall, J. G. & Morris, B. The immediate effect of antigens on the cell output of a lymph node. *Br. J. Exp. Pathol.* **46**, 450–454 (1965).
25. Huang, A. Y., Qi, H. & Germain, R. N. Illuminating the landscape of *in vivo* immunity: insights from dynamic *in situ* imaging of secondary lymphoid tissues. *Immunity* **21**, 331–339 (2004).
26. Lindquist, R. L. *et al.* Visualizing dendritic cell networks *in vivo*. *Nature Immunol.* **5**, 1243–1250 (2004).
27. Nagorsen, D., Marincola, F. M. & Panelli, M. C. Cytokine and chemokine expression profiles of maturing dendritic cells using multiprotein platform arrays. *Cytokine* **25**, 31–35 (2004).
28. Guan, E., Wang, J. & Norcross, M. A. Identification of human macrophage inflammatory proteins 1α and 1β as a native secreted heterodimer. *J. Biol. Chem.* **276**, 12404–12409 (2001).
29. Sherry, B. *et al.* Resolution of the two components of macrophage inflammatory protein 1, and cloning and characterization of one of those components, macrophage inflammatory protein 1β. *J. Exp. Med.* **168**, 2251–2259 (1988).
30. Witt, C. M., Raychaudhuri, S., Schaefer, B., Chakraborty, A. K. & Robey, E. A. Directed migration of positively selected thymocytes visualized in real time. *PLoS Biol.* **3**, e160 (2005).
31. Hommel, M. & Kyewski, B. Dynamic changes during the immune response in T cell-antigen-presenting cell clusters isolated from lymph nodes. *J. Exp. Med.* **197**, 269–280 (2003).
32. Molon, B. *et al.* T cell costimulation by chemokine receptors. *Nature Immunol.* **6**, 465–471 (2005).
33. Janssen, E. M. *et al.* CD4⁺ T cells are required for secondary expansion and memory in CD8⁺ T lymphocytes. *Nature* **421**, 852–856 (2003).
34. Shedlock, D. J. & Shen, H. Requirement for CD4 T cell help in generating functional CD8 T cell memory. *Science* **300**, 337–339 (2003).
35. Banchereau, J. & Steinman, R. M. Dendritic cells and the control of immunity. *Nature* **392**, 245–252 (1998).
36. Catron, D. M., Itano, A. A., Pape, K. A., Mueller, D. L. & Jenkins, M. K. Visualizing the first 50 hr of the primary immune response to a soluble antigen. *Immunity* **21**, 341–347 (2004).
37. Klinman, D. M. Immunotherapeutic uses of CpG oligodeoxynucleotides. *Nature Rev. Immunol.* **4**, 249–258 (2004).
38. von Andrian, U. H. & Mempel, T. R. Homing and cellular traffic in lymph nodes. *Nature Rev. Immunol.* **3**, 867–878 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Reis e Sousa and H. Qi for review of the manuscript, and J. Zhu for advice and assistance with the Q-PCR analyses. This work was supported by the Intramural Research Program of the NIH, NIAID. A.Y.H. is supported by a postdoctoral fellowship grant from the Cancer Research Institute.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.N.G. (rgermain@nih.gov).

LETTERS

Vega is a rapidly rotating star

D. M. Peterson¹, C. A. Hummel^{2,3}, T. A. Pauls⁴, J. T. Armstrong⁴, J. A. Benson⁵, G. C. Gilbreath⁴, R. B. Hindsley⁴, D. J. Hutter⁵, K. J. Johnston³, D. Mozurkewich⁶ & H. R. Schmitt^{4,7}

Vega, the second brightest star in the northern hemisphere, serves as a primary spectral type standard¹. Although its spectrum is dominated by broad hydrogen lines, the narrower lines of the heavy elements suggested slow to moderate rotation, giving confidence that the ground-based calibration of its visible spectrum could be safely extrapolated into the ultraviolet and near-infrared (through atmosphere models²), where it also serves as the primary photometric calibrator. But there have been problems: the star is too bright compared to its peers³ and it has unusually shaped absorption line profiles, leading some^{4,5} to suggest that it is a distorted, rapidly rotating star seen pole-on. Here we report optical interferometric observations that show that Vega has the asymmetric brightness distribution of the bright, slightly offset polar axis of a star rotating at 93 per cent of its breakup speed. In addition to explaining the unusual brightness and line shape peculiarities, this result leads to the prediction of an excess of near-infrared emission compared to the visible, in agreement with observations^{6,7}. The large temperature differences predicted across its surface call into question composition determinations, adding uncertainty to Vega's age and opening the possibility that its debris disk⁸ could be substantially older than previously thought^{9,10}.

Single baseline Michelson stellar interferometers measure complex 'visibilities'¹¹, usually recorded as amplitudes and phases, which are related to the intensity distribution of the target through a Fourier transform. Even though the phases have been shown to be very sensitive to asymmetries in the intensity distribution, they are badly corrupted by the atmosphere and have been little used in the optical. But for closure phases, the data we focus on here—which are obtained by summing the phases measured on each baseline of a triangle in an interferometric array such as the Navy Prototype Optical Interferometer¹² (NPOI)—the atmospheric contribution cancels. The use of closure phase in the radio¹³ has enabled a dramatic gain in dynamic range of interferometric images made from multi-antenna arrays. In the optical^{12,14}, where the phase errors can reach 100 waves on long baselines, the technique enables the use of phase information in any guise. As the observations reported here were made just with a three-telescope array, the application of imaging techniques was not justified and we relied instead on fitting Roche models to the closure phase data.

The application of Roche spheroids to rotating stars was worked out 80 years ago¹⁵. Assuming solid body rotation and a point mass gravitational potential, a rotating star will adopt the figure of a Roche spheroid. Conservation of energy through surfaces of constant potential leads to the prediction that when the energy is transported by radiation the amount transported will vary over the surface in proportion to the effective gravity¹⁶ (that is, gravity minus local centrifugal terms). Near breakup, the effective gravity near the equator can become quite small, leading to the prediction of a

large drop in the local temperature with a corresponding decrease in brightness, an effect referred to as 'gravity darkening'. Rapidly rotating stars seen at intermediate inclinations are therefore expected to display asymmetric intensity distributions. Altair^{17,18} proved to be the first major test of this theory in an isolated rotating star, where the theory succeeded to a high degree in describing a very non-trivial brightness distribution.

The observations considered here were obtained during late May and early June 2001 on the same nights as those of Altair previously reported^{17,19}, where Vega served as a check star (the Vega data in machine readable form are in a separate file in the Supplementary Information). The observations and much of the data reductions are exactly as described for Altair¹⁷, which should be consulted for details. Issues specific to the Vega data are described in the

Table 1 | Vega model and derived parameters

Quantity	Value	Errors (s.d.)*
$\omega = \Omega/\Omega_B^\dagger$	0.926	± 0.021
θ_p (mas) [†]	2.767	0.037
T_p (K) [†]	9,988	61
i (deg.) [†]	4.54	0.33
PA (deg.) [†]	8.6	2.7
v_{eq} (km s ⁻¹)	274	14
$v_{eq,B}^\ddagger$ (km s ⁻¹)	356.1	2.4
Ω (d ⁻¹)	1.884	0.081
$\Omega_B^\ddagger$ (d ⁻¹)	2.034	0.041
T_{eq} (K)	7,557	261
R_p ($R_\odot$)	2.306	0.031
R_{eq} ($R_\odot$)	2.873	0.026
θ_{min}^S (mas)	3.441	0.031
θ_{max}^S (mas)	3.446	0.031
$\log L$ ($L_\odot$)	1.544	0.018
$\log g_p$ (cm ² s ⁻²)	4.074	0.012
$\log g_{eq}$ (cm ² s ⁻²)	3.589	0.056
M ($M_\odot$)	2.303	0.024
$T_{eff }$ (K)	9,306	86
Age $_{ }$ (Myr)	386	16

Six quantities are needed to uniquely define the Roche model of a star¹⁷: the ratio of the angular rotation to that of breakup, $\omega = \Omega/\Omega_B$, the inclination (or tilt) of the rotational axis to the line of sight, i , the position angle, PA, of the pole on the sky, the radius of the polar axis, R_p (or equivalently, using the parallax, the polar angular diameter, θ_p), the effective temperature at the pole, T_p , and the surface gravity at the pole, g_p , or equivalently, the mass. It is then possible to calculate the radius, $R(\phi)$, of the star for a given stellar latitude, ϕ , the effective gravity and the local temperature ($T(\phi)$). For the spectral calculations we have adopted the ATLAS model atmospheres²⁸ and in particular the Van Hamme²⁹ limb-darkening parameterization of that grid. Other parameters include linear rotational velocities, v , the luminosity, L , surface gravities, g , the mass, M , and effective temperature, T_{eff} , the last referring to the entire non-rotating star. Subscripts eq and p specify quantities evaluated at the equator and pole, respectively.

*Uncertainties due to the parallax have not been included in the errors.

[†]The parameters derived from the model fit. A mass of 2.30 $M_\odot$ was assumed in the fit.

[‡]Rotating at breakup but with the same mass and polar radius.

$\theta_{min,max}^S$ are the minimum and maximum projected angular diameters.

$||$ The parameters of a non-rotating star from the Padova grid²⁶ that would reproduce the (corrected) luminosity and polar radius¹⁷.

¹Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794-3800, USA. ²European Southern Observatory (ESO), Casilla 19001, Santiago 19, Chile. ³US Naval Observatory, 3450 Massachusetts Avenue NW, Washington DC, 20392-5420, USA. ⁴Naval Research Laboratory, Code 7215, 4555 Overlook Avenue SW, Washington DC 20375, USA. ⁵US Naval Observatory, Flagstaff Station, 10391 West Naval Observatory Road, Flagstaff, Arizona 86001-8521, USA. ⁶Seabrook Engineering, 9310 Dubarry Road, Seabrook, Maryland 20706, USA. ⁷Interferometrics, Inc., 13454 Sunrise Valley Drive, Suite 240, Herndon, Virginia 20171, USA.

Table 2 | Predicted photometric ‘excesses’ for Vega

Band	Wavelength (μm)	Excess (mag)
R	0.67	−0.016
I	0.86	−0.029
J	1.12	−0.052
K	2.14	−0.072
L	3.69	−0.072

The differential excesses for the model of Vega compared to a non-rotating model (both solar composition) which matches the V magnitude with an angular diameter of 3.24 mas (ref. 27). Until the question of composition is resolved, these values should be considered indicative.

Supplementary Information. We focus here on the closure phase data taken by NPOI on 25 May 2001, as they are the most extensive and the highest quality data of that run. We augment these with the V-band magnitude², $V = 0.026 \pm 0.008$, as an additional observable. The model was fitted by enforcing the usual minimum χ^2 metric using the Levenberg–Marquardt algorithm²⁰. The parameters from the initial reduction are given in column 2 of Supplementary Table S1. The projected (at inclination i) equatorial velocity, v_{eq} , for this model was predicted to be $v_{\text{eq}} \sin i \approx 15 \text{ km s}^{-1}$, a bit below the $\sim 21.8 \pm 0.1 \text{ km s}^{-1}$ found from detailed profile fits using a rotating model²¹. Although it is not so far off given how close the star is to pole-on, we feel the projected velocity is sufficiently well

known that it should also be incorporated in the fit. We have therefore added as an ‘observable’ $v_{\text{eq}} \sin i = 22 \pm 1.0 \text{ km s}^{-1}$ (see Supplementary Information) and refit the data (columns 3 and 4 of Supplementary Table S1). The Roche model provides a good fit to the augmented data set. The model parameters fitted, a number of derived quantities, and an estimate of what the main parameters of the star Vega would be, were it not rotating, are given in Table 1. The quality of the fit is illustrated in Fig. 1, and a false-colour rendering of the Vega model is shown in Fig. 2. As Vega has so many roles in astronomy, the ramifications of this result are extensive. We summarize some of the most important below. Rotation can affect the gross spectral distribution of a star, an issue of considerable import given Vega’s role as the primary flux calibrator in the ultraviolet, visible and near-infrared. To estimate the size of these effects, we have calculated the changes in the fluxes from the rotating model compared to the static case that would be measured through a series of broadband filters (Table 2). As can be seen, there is a significant, systematic increase in the infrared emission from the rotating model. This is understood as being due to the large amount of surface area predicted to be at relatively low temperatures. There is an extensive literature on the possibility and extent of an ‘infrared excess’ in the Vega spectrum^{6,7}. As the issue is critical to so much of astronomy, sides have been strongly taken. Observations of

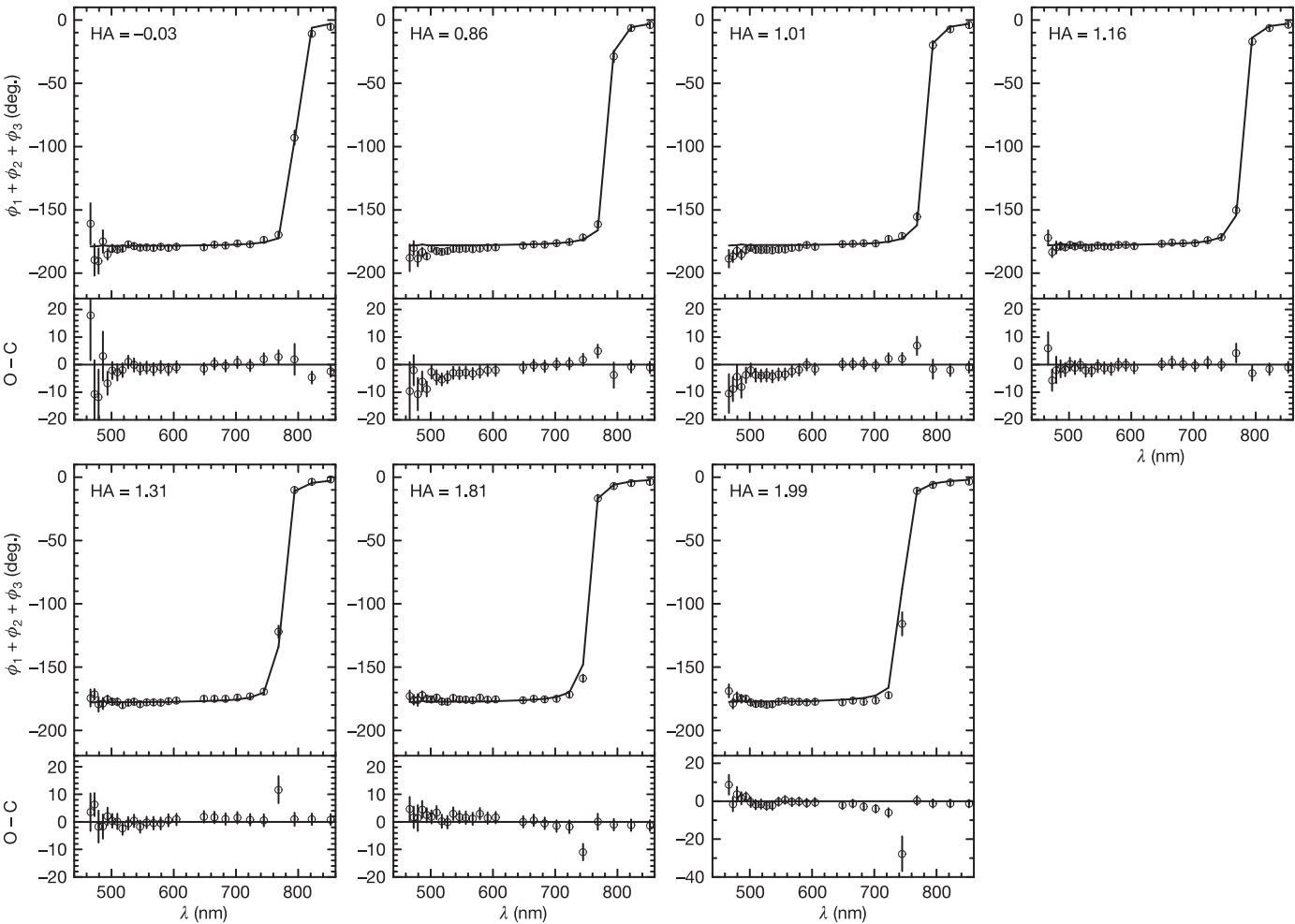


Figure 1 | Roche model fits to the closure phase data taken on 25 May 2001. The observations (open circles), estimated errors (bars, standard deviations) and model calculations (solid lines) are shown for each scan (labelled by hour angle, HA). Residuals (observed, O, minus calculated, C) are shown below each of the scans for clarity. The phases for the individual baselines, ϕ_b , and thus the closure phases, take on only two values, 0° or 180° ,

if an object is centro-symmetric. Closure phase measurements showing departures from this simple ‘abrupt transition’ behaviour provide potentially very sensitive measurements of asymmetry in an object. The soft transition at the points of the 180° phase changes here give a clear signal of the asymmetry in the intensity distribution. (Note the scale change for the residuals of the last scan.)

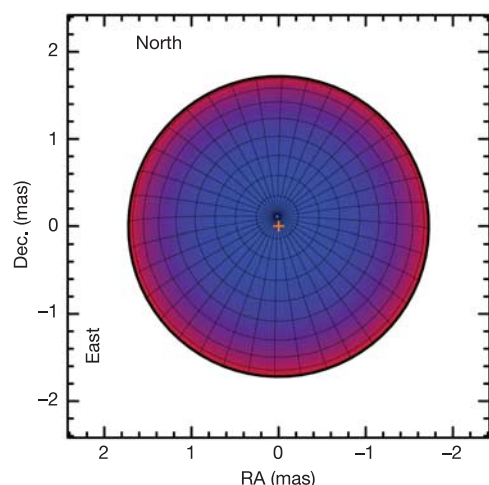


Figure 2 | A false-colour model of Vega as it appears from Earth. (Blue is bright, red is faint, and the orange '+' is the subsolar point.) The temperature drops more than 2,400 K from pole to equator, creating an 18× drop in intensity at 500 nm. Limb-darkening in a non-rotating model predicts only a 5-fold drop in intensity. Although the projected outline is almost perfectly circular, the polar diameter is only 80% of the equator. Dec., declination; RA, right ascension.

other A stars showed that Vega's colours appear sensibly normal⁷, which led some authors^{7,22} to wonder whether the problem was with the model atmospheres. Others²³ argued that, because the hydrogen absorption coefficient completely dominates the spectra of A stars and is so well known, it was unlikely that the model atmospheres were wrong, an argument that seems to have carried the day. Rapid rotation provides a simple resolution to this controversy—Vega is best modelled as a composite of model atmospheres. The star should have 'normal' infrared colours, as most normal A stars are rapid rotators, and there is no need to question the quality of the individual atmosphere models.

One important aspect to this model is the near pole-on orientation of the rotational axis. Vega is surrounded by a large infrared-emitting disk of material, a 'debris disk', which presents an essentially circular profile. We do not expect perfect coupling between orbiting material and the central star. But, if the poles of the disk and Vega coincide and the disk is thin, we would predict 0.3% flattening, which is unlikely to be detectable, as seems to be the case²⁴.

Related to the orientation of the pole is the inferred equatorial velocity, $v_{\text{eq}} = 272 \text{ km s}^{-1}$ (Table 1). Among the Vega-like stars, Vega itself has been anomalous in displaying a very low projected rotational velocity¹⁰. The present determination strongly supports the view that the Vega-like stars are rapid rotators, consistent with the large amounts of angular momentum in the surrounding dust clouds.

Vega's rotational state also affects inferences about its debris disk. It is well known that rotation results in the apparent brightness, and in turn the deduced luminosity, being inclination-dependent²⁵. Song *et al.*¹⁰, for example, have gone to some lengths to characterize the uncertainty introduced by this effect in the Vega-like stars. It is straightforward to calculate the total luminosity of a Roche spheroid (Table 1). Further, one can apply small corrections¹⁷ to the luminosity and polar radius to obtain the corresponding values that would apply to a non-rotating star of the same mass. Using the Padova²⁶ models, we derive a mass $M = 2.303 M_{\odot}$ (where $M_{\odot}$ is the solar mass) and an age of $386 \pm 16 \text{ Myr}$, on the high side of recent estimates of $354^{+29}_{-87} \text{ Myr}$ (ref. 10) and $347^{+43}_{-37} \text{ Myr}$ (ref. 9).

Unfortunately, Vega's composition enters the age determination rather critically. The star is currently viewed as underabundant in heavy elements compared to the Sun²⁷, $[\text{Fe}/\text{H}] \approx -0.5$, which clearly needs to be revisited, given the 2,400 K temperature drop

now predicted across its surface. Previous age determinations have implicitly or explicitly assumed solar metallicity, the argument being that compositional peculiarities among the slowly rotating stars in this part of the Hertzsprung–Russell diagram are probably limited to the outer envelope, and normal composition models are therefore appropriate for estimating bulk properties. In our interpretation that argument fails, because at these rotational velocities meridional circulation will keep the bulk of the star well mixed.

Proceeding as above, we estimate $M = 2.11 M_{\odot}$ and an age of 572 Myr, using the metallicity $Z = 0.008$ ($[\text{Fe}/\text{H}] \approx -0.4$) Padova models²⁶. It is clear that a full abundance analysis incorporating rotation needs to be performed to remove this uncertainty. In the meantime this range, 386–572 Myr, is probably a more realistic estimate of the uncertainty in the evolutionary age of Vega and its debris disk.

Received 10 January; accepted 14 February 2006.

- Morgan, W. W. & Keenan, P. C. Spectral classification. *Annu. Rev. Astron. Astrophys.* **11**, 29–50 (1973).
- Bohlin, R. C. & Gilliland, R. L. Hubble Space Telescope absolute spectrophotometry of Vega from the far-ultraviolet to the infrared. *Astron. J.* **127**, 3508–3515 (2004).
- Petrie, R. M. Spectroscopic absolute magnitudes of early-type stars from hydrogen absorption—a revision. *Publ. Dominion Astrophys. Obs.* **12**, 317–338 (1964).
- Gray, R. O. The spectroscopic and photometric effects of rotation in the A-type stars. *J. R. Astron. Soc. Can.* **82**, 336–348 (1988).
- Gulliver, A. F., Hill, G. & Adelman, S. J. Vega: A rapidly rotating pole-on star. *Astrophys. J.* **429**, L81–L84 (1994).
- Blackwell, D. E., Leggett, S. K., Petford, A. D., Mountain, C. M. & Selby, M. J. Absolute calibration of the infrared flux from Vega at 1.24, 2.20, 3.76 and 4.6 microns by comparison with a standard furnace. *Mon. Not. R. Astron. Soc.* **205**, 897–905 (1983).
- Leggett, S. K., Bartholomew, M., Mountain, C. M. & Selby, M. J. Narrow-band 1- to 5- μm photometry of A-type stars. *Mon. Not. R. Astron. Soc.* **223**, 443–447 (1986).
- Aumann, H. H. *et al.* Discovery of a shell around Alpha Lyrae. *Astrophys. J.* **278**, L23–L27 (1984).
- Habing, H. J. *et al.* Disappearance of stellar debris disks around main-sequence stars after 400 million years. *Nature* **401**, 456–458 (1999).
- Song, I., Caillault, J.-P., y Navascués, D. B. & Stauffer, J. R. Ages of A-type Vega-like stars from uvby β photometry. *Astrophys. J.* **546**, 352–357 (2001).
- Thompson, R. A., Moran, J. & Swenson, G. Jr *Interferometry and Synthesis in Radio Astronomy* (Wiley-VCH, Weinheim, 2004).
- Armstrong, J. T. *et al.* The Navy Prototype Optical Interferometer. *Astrophys. J.* **496**, 550–571 (1998).
- Jennison, R. C. A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent. *Mon. Not. R. Astron. Soc.* **118**, 276–284 (1958).
- Baldwin, J. E. *et al.* The first images from an optical aperture synthesis array: mapping of Capella with COAST at two epochs. *Astron. Astrophys.* **306**, L13–L16 (1996).
- von Zeipel, H. The radiative equilibrium of a slightly oblate rotating star. *Mon. Not. R. Astron. Soc.* **84**, 684–701 (1924).
- Hardorp, J. & Strittmatter, P. A. The effect of rotation on radiation from stars I. Properties of the continuum. *Astrophys. J.* **151**, 1057–1073 (1968).
- Peterson, D. M. *et al.* Resolving the effects of rotation in Altair with long-baseline interferometry. *Astrophys. J.* **636**, 1087–1097 (2006).
- Domiciano de Souza, A. *et al.* Gravitational-darkening of Altair from interferometry. *Astron. Astrophys.* **442**, 567–578 (2005).
- Ohishi, N., Nordgren, T. E. & Hutter, D. J. Asymmetric surface brightness distribution of Altair observed with the Navy Prototype Optical Interferometer. *Astrophys. J.* **612**, 463–471 (2004).
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. *Numerical Recipes in C* 2nd edn (Cambridge Univ. Press, Cambridge, 1992).
- Hill, G., Gulliver, A. F. & Adelman, S. J. in *The A-Star Puzzle* (eds Zverko, J., Ziznovsky, J., Adelman, S. J. & Weiss, W. W.) 35–42 (IAU Symp. 224, Cambridge Univ. Press, Cambridge, 2004).
- Leggett, S. K. The flux distribution of Vega for the wavelength region from 10 to 100 microns, and the calibration of IRAS at 12 microns and 25 microns. *Astron. Astrophys.* **153**, 273–277 (1985).
- Bessel, M. S., Castelli, F. & Plez, B. Model atmospheres broad-band colors, bolometric corrections and temperature calibrations for O – M stars. *Astron. Astrophys.* **333**, 231–250 (1998).
- Su, K. Y. L. *et al.* The Vega debris disk: A surprise from Spitzer. *Astrophys. J.* **628**, 487–500 (2005).
- Collins, G. W. II Continuum emission from a rapidly rotating stellar atmosphere. *Astrophys. J.* **138**, 1134–1146 (1963).

26. Girardi, L., Bressan, A., Bertelli, G. & Chiosi, C. Evolutionary tracks and isochrones for low- and intermediate-mass stars: From 0.15 to 7 $M_{\odot}$, and from $z = 0.0004$ to 0.03. *Astron. Astrophys. Suppl.* **141**, 371–383 (2000).
27. Castelli, F. & Kurucz, R. L. Model atmospheres for Vega. *Astron. Astrophys.* **281**, 817–832 (1994).
28. Kurucz, R. L. *ATLAS9 Stellar Atmospheres and 2 km/s Grid* (CD-ROM No. 13, Smithsonian Astrophysical Observatory, Cambridge, MA, 1993); (<http://kurucz.harvard.edu/cdroms.html>).
29. Van Hamme, W. New limb-darkening coefficients for modeling binary light curves. *Astron. J.* **106**, 2096–2117 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Kurucz for advice on calculating broadband fluxes using model atmospheres. The NPOI facility is a collaboration between the Naval Research Laboratory and the US Naval Observatory in association with Lowell Observatory, and is funded by the Office of Naval Research and the Oceanographer of the Navy. This research has made use of the SIMBAD literature database, operated at CDS, Strasbourg, France, and of NASA's Astrophysics Data System.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.M.P. (dpeterson@astro.sunysb.edu).

LETTERS

A quantum Newton's cradle

Toshiya Kinoshita¹, Trevor Wenger¹ & David S. Weiss¹

It is a fundamental assumption of statistical mechanics that a closed system with many degrees of freedom ergodically samples all equal energy points in phase space. To understand the limits of this assumption, it is important to find and study systems that are not ergodic, and thus do not reach thermal equilibrium. A few complex systems have been proposed that are expected not to thermalize because their dynamics are integrable^{1,2}. Some nearly integrable systems of many particles have been studied numerically, and shown not to ergodically sample phase space³. However, there has been no experimental demonstration of such a system with many degrees of freedom that does not approach thermal equilibrium. Here we report the preparation of out-of-equilibrium arrays of trapped one-dimensional (1D) Bose gases, each containing from 40 to 250 ⁸⁷Rb atoms, which do not noticeably equilibrate even after thousands of collisions. Our results are probably explainable by the well-known fact that a homogeneous 1D Bose gas with point-like collisional interactions is integrable. Until now, however, the time evolution of out-of-equilibrium 1D Bose gases has been a theoretically unsettled issue^{4–6}, as practical factors such as harmonic trapping and imperfectly point-like interactions may compromise integrability. The absence of damping in 1D Bose gases may lead to potential applications in force sensing and atom interferometry.

To see qualitatively why 1D gases might not thermalize, consider the elastic collision of two isolated, identical mass classical particles in one dimension. Energy and momentum are conserved only if they simply exchange momenta. Clearly, the momentum distribution of a 1D ensemble of particles will not be altered by such pairwise collisions. The well-known behaviour of Newton's cradle (see Fig. 1a) is most easily understood in this way. Even when several balls are simultaneously in contact, particles in an idealized Newton's cradle just exchange specific momentum values, though the explanation is more subtle⁷. Generalization of the Newton's cradle to quantum mechanical particles lends it a ghostly air. Rather than just reflecting off each other, colliding particles can also transmit through each other. When the particles are identical, the final states after transmission and reflection are indistinguishable.

In general, correlations and overlap among 1D Bose gas wavefunctions complicate the picture of independent particles colliding as in a Newton's cradle. In fact, there are circumstances in which 1D momentum distributions are known to change in time. For example, when weakly coupled bosons are released from a trap, the conversion of mean field energy to kinetic energy changes the momentum distribution. In the Tonks–Girardeau limit of infinite strength interactions⁸, although the 1D bosons interact locally like non-interacting fermions, their momentum distribution is not fermionic^{9,10}. When a Tonks–Girardeau gas is released from a trap and expands in one dimension, its momentum distribution evolves into that of a trapped Fermi gas^{11–13}. The quantum Newton's cradle view of particles colliding with each other and either reflecting or transmitting can only be applied when the kinetic energy of the collision greatly exceeds the energy per atom at zero temperature at

the prevailing density¹⁴. The collisions that we study satisfy this criterion well. Our observations extend from the Tonks–Girardeau regime, where only pairwise collisions can occur¹⁵, to the intermediate coupling regime, where there can be three- (or more) body collisions^{15–17}. In both regimes, atoms that are set oscillating and colliding in a trap do not appreciably thermalize during our experiment.

We start our experiments with a Bose–Einstein condensate (BEC) loaded into the combination of a blue-detuned two-dimensional (2D) optical lattice and a red-detuned crossed dipole trap (see Methods). The combination of light traps makes a 2D array of distinct, parallel Bose gases, with the 2D lattice providing tight transverse confinement and the crossed dipole trap providing weak axial trapping¹¹. The dynamics within each tube of the 2D array are strictly 1D because the lowest transverse excitation, $\hbar\omega_r$ (where $\omega_r/2\pi = 67$ kHz is the transverse oscillation frequency), far exceeds all other energies in

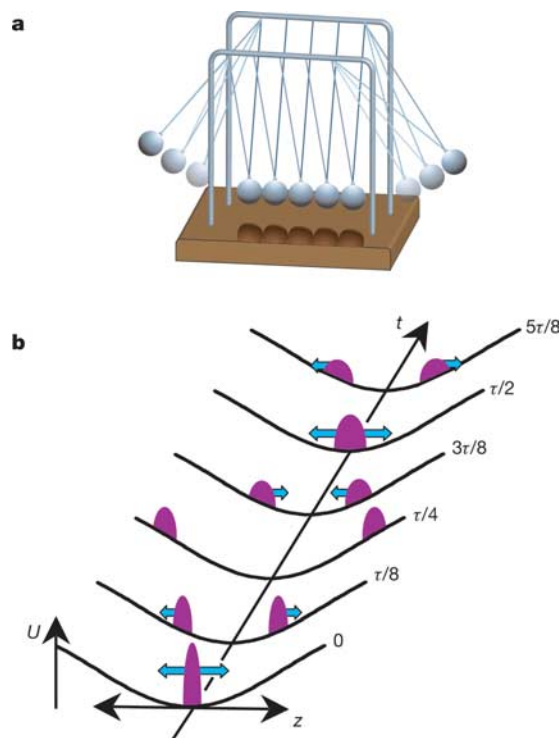


Figure 1 | Classical and quantum Newton's cradles. **a**, Diagram of a classical Newton's cradle. **b**, Sketches at various times of two out of equilibrium clouds of atoms in a 1D anharmonic trap, $U(z)$. At time $t = 0$, the atoms are put into a momentum superposition with $2\hbar k$ to the right and $2\hbar k$ to the left. The two parts of the wavefunction oscillate out of phase with each other with a period τ . Each atom collides with the opposite momentum group twice every full cycle, for instance, at $t = 0$ and $\tau/2$. Anharmonicity causes each group to gradually expand, until ultimately the atoms have fully dephased. Even after dephasing, each atom still collides with half the other atoms twice each cycle.

¹Department of Physics, The Pennsylvania State University, 104 Davey Laboratory, University Park, Pennsylvania 16802, USA.

the problem, and there is negligible tunnelling among the tubes. We can vary the weighted average number of atoms per tube, N_{tube} , and the axial oscillation period, τ . For a given array, τ is the same to within 6% for all 1,000–8,000 tubes. The 1D coupling strength is given by $\gamma = |2/a_{1D}n_{1D}|$, where n_{1D} is the 1D density, $|a_{1D}| \approx a^2/2a$ is the 1D scattering length, $a = 5.3$ nm is the three-dimensional (3D) scattering length, $a_r = (\hbar/m\omega_r)^{1/2} = 41.5$ nm is the transverse oscillator width, and m is the Rb mass¹⁸.

To study the 1D Bose gases, we turn off the crossed dipole trap and allow the atoms to expand in one dimension for 27 ms before taking an absorption image from the transverse direction. When we integrate the image transverse to the tubes, we get a 1D spatial distribution that corresponds to the momentum distribution after expansion, $f(p_{\text{ex}})$. Although the individual 1D gases have Thomas–Fermi or Tonks–Girardeau $f(p_{\text{ex}})$ profiles, we measure gaussian $f(p_{\text{ex}})$ distributions, as expected when the $f(p_{\text{ex}})$ for many 1D Bose gases with different N_{tube} are summed.

To create non-equilibrium momentum distributions, we pulse on a 3.2 THz detuned 1D lattice along the tubes, which acts as a phase grating for the atoms. Two pulses, with intensity 11 W cm^{-2} and pulse widths of $23 \mu\text{s}$ separated in time by $33 \mu\text{s}$, can deplete the

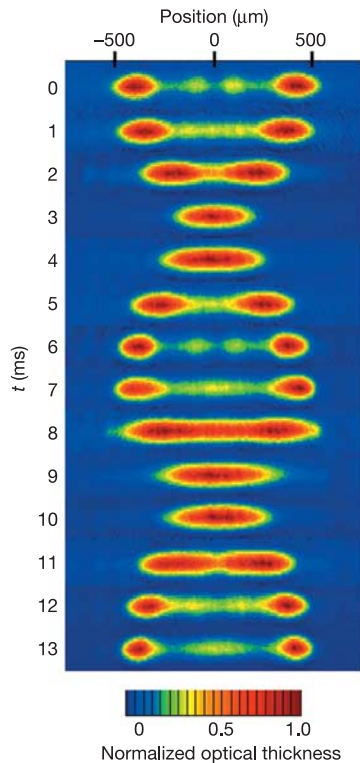


Figure 2 | Absorption images in the first oscillation cycle for initial average peak coupling strength $\gamma_o = 1$. Atoms are always confined to one dimension, in this case in 3,000 parallel tubes, with a weighted average of 110 atoms per tube. After grating pulses put each atom in a superposition of $\pm 2\hbar k$ momentum, they are allowed to evolve for a variable time t in the anharmonic 1D trap (crossed dipole trap), before being released and photographed 27 ms later. The false colour in each image is rescaled to show detail. These pictures are used to determine $f(p_{\text{ex}})$. The first image shows that some atoms remain near $p_{\text{ex}} = 0$ at $t = 0$. How many remain there depends on n_{1D} , implying that these remnant atoms do not result from an imperfect pulse sequence, but rather from interactions during the grating pulses or evolution of the momentum distribution during expansion. The relative narrowness of the peaks in the last image compared to the first is indicative of the reduction in spatial density that results from dephasing (Fig. 1b). The transverse spatial width of each of the 14 image frames is $70 \mu\text{m}$. Horizontal in the figure corresponds to vertical in the experiment, a minor distinction because a magnetic field gradient cancels gravity for the atoms.

zero momentum state and transfer atoms to $\pm 2\hbar k$ peaks^{19,20} where k is the wavevector of the 1D lattice light. We wait after the grating pulses for a variable time, t , before measuring $f(p_{\text{ex}})$. Figure 2 shows a time series of absorption images spanning a full oscillation in the crossed dipole trap, when the weighted average of the initial peak γ in each tube, γ_o , is 1.0. The two momentum groups collide with each other in the centre of the crossed dipole trap twice each full cycle, for instance at $t = 0$ and $\tau/2$, as illustrated in Fig. 1b. The total collision energy is $8(\hbar k)^2/2m = 0.45\hbar\omega_r$, less than one-quarter the energy needed for transverse vibrational excitation²¹, so the colliding gases remain 1D.

The first and last images in Fig. 2 differ because the oscillating atoms dephase. Illustrated conceptually in Fig. 1b, there is dephasing due to the gaussian crossed dipole trap anharmonicity, which gives an $\sim 8\%$ spread of τ across the full-width at half-maximum of each of the colliding clouds. The top curves in Fig. 3a–c show the time-averaged $f(p_{\text{ex}})$ over the first cycle for different γ_o . Differences in shape among them reflect the initial energy per particle, which increases with n_{1D} , and hence γ_o^{-1} . Within 10τ to 15τ , $f(p_{\text{ex}})$ stops changing noticeably during an oscillation period. The central observations in this letter are of the evolution of $f(p_{\text{ex}})$ that are dephased, like the lower curves of Fig. 3a–c. Comparing only dephased distributions avoids the complication of how the momentum distribution in the trap evolves into $f(p_{\text{ex}})$ during expansion, which may slightly depend on the initial spatial distributions. As atoms have clearly dephased within each tube, dephasing among tubes is irrelevant.

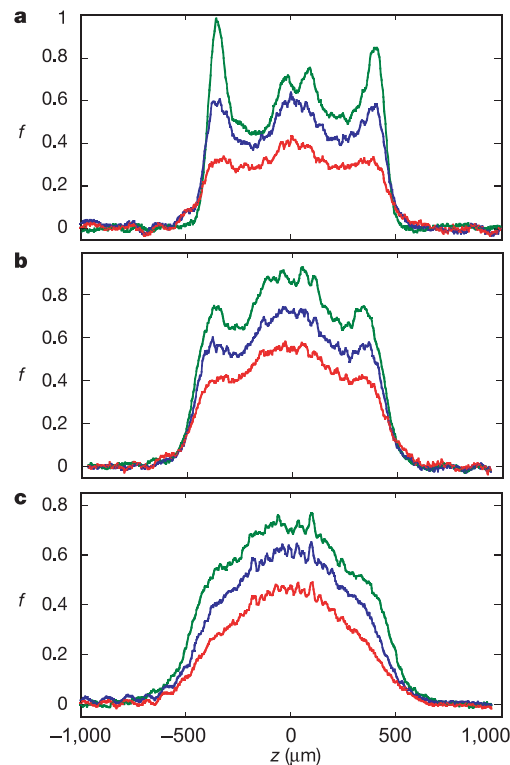


Figure 3 | The expanded momentum distribution, $f(p_{\text{ex}})$, for three values of γ_o . The curves are obtained by transversely integrating absorption images like those in Fig. 2. The spatial position, z , is approximately proportional to the expanded momentum, p_{ex} . The vertical scale is arbitrary, but consistent among the curves. **a**, $\gamma_o = 4$; **b**, $\gamma_o = 1$; and **c**, $\gamma_o = 0.62$. The highest (green) curve in each set is the average of $f(p_{\text{ex}})$ from the first cycle, that is, from the images like those in Fig. 2. The lower curves in each set are $f(p_{\text{ex}})$ taken at single times, t , after the atoms have dephased: **a**, $\tau = 34$ ms, $t = 15\tau$ (blue) and 30τ (red); **b**, $\tau = 13$ ms, $t = 15\tau$ (blue) and 40τ (red); and **c**, $\tau = 13$ ms, $t = 15\tau$ (blue) and 40τ (red). The changes in the distribution with time are attributable to known loss and heating. (See Supplementary Information for a discussion of the fine spatial structure in these curves.)

All the curves in Fig. 3 are non-gaussian. For comparison, we have created equilibrium 1D Bose gases with the same r.m.s. momentum as the non-equilibrium distributions we study here. To do so, we start with an equilibrium 3D Bose gas at an elevated temperature and adiabatically turn on the 2D lattice. The resultant $f(p_{\text{ex}})$ are nearly perfectly gaussian. Thus, to the extent that an observed $f(p_{\text{ex}})$ is not gaussian, it has not thermalized.

Heating and loss affect the evolution of the distribution. We have studied these processes by watching how $f(p_{\text{ex}})$ evolves without any grating pulses (see Supplementary Information). Some loss (20% or less, depending on γ_o) comes in the first couple of hundred milliseconds from three-body inelastic collisions. There is also 15% per second loss to background gas collisions. Spontaneous emission caused by the lattice light heats some atoms, and by leaving some atoms in unlevitated magnetic sublevels, causes a 30% per second loss. This last loss in turn causes most of the heating, as exiting atoms transfer some of the momentum they pick up on their way out to atoms that remain.

To account for loss and heating in the time evolution shown in Fig. 3, we project how already dephased distributions would evolve

without thermalization. Specifically, we take $f(p_{\text{ex}})$ at a time $t_o = 15\tau$, rescale it to account for loss during an observation time, t_{obs} , and convolve it with gaussian widths to capture the effect of the independently measured heating during t_{obs} (see Supplementary Information). The blue curves in Fig. 4 were projected with a two-component model that accurately reflects the measured heating, for $\gamma_o(\gamma_d) = 4(18)$, 1(3.2) and 0.62(1.4), where the coupling strength after dephasing, γ_d , is calculated using the reduced n_{1D} that prevails at t_o . The green curves are the result of a simpler single-component projection. The similarity of the blue and green lines illustrates the robustness of our projections (see Supplementary Information). The red curves show the actual distributions after t_{obs} .

The actual and projected curves overlap reasonably well, with reduced χ^2 values of 1.2, 1.35 and 2.5 for Fig. 4a, b and c, respectively (using the blue curves). In each case, the difference between the projected and actual curves is far smaller than the difference between either of them and a thermal distribution. To highlight the non-gaussian shape of Fig. 4c, we have superimposed a gaussian with the same atom number and r.m.s. width as the data. The slight discrepancies that exist between the actual and projected curves may result from the $\sim 25\%$ loss of atoms during t_{obs} , which reduces the interaction energy contribution to $f(p_{\text{ex}})$. By assuming that any deviation between the projected and actual distributions is a step along the way to thermalization, we conservatively determine a lower bound on the thermalization time constant, τ_{th} (see Methods). τ_{th} is at least 390τ , $1,910\tau$ and 200τ for $\gamma_d = 18$, 3.2 and 1.4, respectively. The data imply that each atom continues to oscillate in the trap with the same peak momentum it was given initially, as if there were no collisions.

Although collisions have no dynamical effect, we would like to roughly keep track of how many have occurred. Each atom passes $N_{\text{tube}}/2$ atoms every half cycle. The probability of reflection, R , in a pairwise collision of 1D bosons with centre of mass momentum $2\hbar k$ was calculated in ref. 22. In the limit where $(2ka_{1D})^2 \gg 1$, $R = (2ka_{1D})^{-2}$. For our confinement parameters, $R = 1/22$. Therefore, in the first full cycle, the number of $2\hbar k$ collisions is N_{tube} , with $r = N_{\text{tube}}/22$ reflections. After dephasing within a tube, each atom has as many collisions, but at centre of mass momenta that range from $2\hbar k$ to near 0. As the relative velocity decreases, R increases quadratically (until it saturates), but the ability of a collision to redistribute momentum is reduced roughly quadratically. Accordingly, we use the r derived above to keep track of reflections even after the atoms have dephased. For the conditions in Fig. 4a, b and c, the average number of collisions that have occurred per atom during t_{obs} are 600, 2,750 and 6,250, respectively, and the average number of reflections are 27, 125 and 285. Using the results from Fig. 4, we can set lower limits on the number of reflections required for thermalization of 710, 9,600 and 2,300 for $\gamma_d = 18$, 3.2 and 1.4, respectively. These limits are obviously much larger than the 2.7 collisions that characterize thermalization in a 3D gas²³.

To experimentally confirm the existence of collisions in this system, despite their lack of consequence in one dimension, we apply the grating pulses without ever having turned on the 2D optical lattice, and so create non-equilibrium momentum distributions in three dimensions. Two BECs with different centre of mass velocities collide every half cycle. At the quarter cycle times, the two BECs are well separated spatially. This implies that collisions occur well above the Landau critical velocity, allowing particles to scatter out of the macroscopically occupied states²⁴. We observe thermalization in a two-step process. Atoms first scatter into a spherical shell in velocity, which corresponds to the outgoing s -wave. They then scatter into a broad range of final states. Even though the 3D densities are nearly an order of magnitude lower than in the 1D tubes, thermalization occurs on a $< 2\tau$ timescale.

The absence of damping in 1D Bose gases has several potential applications. Atoms undergoing Bloch oscillations in quantum degenerate gases are candidate force sensors²⁵. Fermions have emerged as better for this purpose than bosons, because the absence

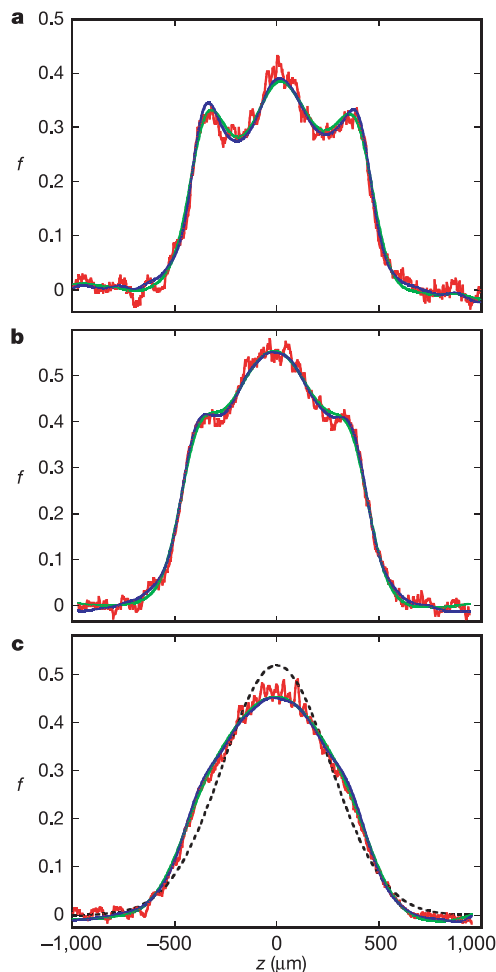


Figure 4 | Projected versus actual $f(p_{\text{ex}})$ for various γ_d , the dephased average peak coupling strength. The blue and green curves are $f(p_{\text{ex}})$ for $t_o = 15\tau$, rescaled to account for loss and convolved with the known heating during t_{obs} . The blue curve's heating model is more sophisticated than that of the green curve, but the results are insensitive to the details. The red curves are the actual distributions at $t_o + t_{\text{obs}}$. **a**, $\gamma_d = 18$ and $t_{\text{obs}} = 15\tau$. **b**, $\gamma_d = 3.2$ and $t_{\text{obs}} = 25\tau$. **c**, $\gamma_d = 1.4$ and $t_{\text{obs}} = 25\tau$. The dashed line in **c** is a gaussian with the same number of atoms and r.m.s. width as the actual distribution. To the extent that the actual distribution conforms to the projected distribution rather than to a gaussian, the atoms have not thermalized.

of s -wave collisions prevents instabilities in the Bloch oscillations²⁶, allowing for longer observation times²⁵. However, the larger fermion momentum spreads reduce device sensitivity. The addition of 1D confinement for bosons could allow for narrow momentum spreads while avoiding damping-related instabilities.

Atom interferometers can be sensitive gyroscopes²⁷, and work is in progress to achieve greater sensitivity using atom waveguides^{20,28}. To date, these experiments have been performed outside the strictly 1D regime. The absence of thermalizing collisions promises potentially large sensitivity gains from reaching one dimension in the Thomas–Fermi regime.

Integrable systems, rare though they are, are central to our understanding of nonlinear dynamics. The KAM theorem¹, which underlies much of nonlinear dynamics, describes the effect of weak non-integrability on system dynamics. Any weak non-integrability in our system is apparently insufficient for thermalization, which confronts theory with a clear and rare many-body experimental observation in this area. In the future, we can allow tunnelling among the tubes by simply lowering the intensity of one or both of the 2D lattice beams, and so continuously change the system from one with 1D to one with 2D or 3D collisional character. The many-body system's integrability can thus be compromised in a controlled way, presenting a new type of experimental model for the onset of chaos. The possibility of direct control of τ_{th} may have further applications in the study of non-equilibrium quantum dynamics²⁹.

In summary, we have watched the time evolution of non-equilibrium trapped 1D Bose gases, which are almost integrable systems. We find no evidence of redistribution of momentum, from the Tonks–Girardeau gas limit to the intermediate coupling regime. That is, we observe thousands of parallel 1D Bose gases, each with hundreds of atoms colliding thousands of times without approaching equilibrium.

METHODS

Trapping details. We adiabatically load an ⁸⁷Rb Bose–Einstein condensate (BEC), produced by all-optical means³⁰ and consisting of 2.5×10^5 atoms, into the 2D lattice/crossed dipole trap combination. A vertical magnetic field gradient cancels the effect of gravity on the $F = 1$, $m_F = 1$ atoms. We vary τ from 13 to 34 ms by changing the power and width of the crossed dipole trap. N_{tube} depends on the value of τ when the 2D lattice is turned on, and ranges from 40 to 250. We vary n_{1D} in a range from 1.7×10^6 to 1.2×10^7 atoms m^{-1} by varying N_{tube} and by varying τ during the observations.

Determination of a lower limit on τ_{th} . To determine a lower limit on τ_{th} , we calculate χ_{ap}^2 , the reduced χ^2 obtained from comparing the actual $f(p_{\text{ex}})$ to the projected $f(p_{\text{ex}})$, and χ_{ag}^2 , the reduced χ^2 obtained from comparing the actual $f(p_{\text{ex}})$ to a gaussian with the same r.m.s. width and number of atoms. We take χ_{ag}^2 as a measure of the extent to which the gas is not thermalized. If we then assume that the entire χ_{ap}^2 represents the small first step in an exponential approach of the actual distribution to a gaussian, then $\tau_{\text{th}} = \tau_{\text{obs}}(\chi_{\text{ag}}^2 + \chi_{\text{ap}}^2)/\chi_{\text{ap}}^2$. Note that even if the agreement between the projected and actual curves was statistically perfect, so that $\chi_{\text{ap}}^2 = 1$, we would still only be able to set a finite lower limit on τ_{th} .

We obtain the variance that normalizes χ_{ap}^2 and χ_{ag}^2 using the tails of the measured $f(p_{\text{ex}})$. Fluctuations in the absorption probe beam dominate the noise, along with the fine spatial structures discussed in Supplementary Information. The spatial frequencies of these noise sources are higher than the spatial frequencies in the projected $f(p_{\text{ex}})$, so they do not contribute to χ_{ap}^2 . But they do reduce χ_{ag}^2 in a physically meaningless way. To get a more appropriate measure of the relative sizes of these two χ^2 , we apply a point spread filter function to the actual $f(p_{\text{ex}})$. We increase the number of points used until the χ_{ap}^2 for the filtered distribution is 10% larger than for the unfiltered function. This is approximately when filtering starts to change the shape of the actual distribution on a scale relevant to the projected distribution. The projected distributions are unaffected by such point spread filtering, as the projection operation smooths those curves. Comparisons of the projected distributions and the filtered actual distributions (similar to Fig. 4) are shown in Supplementary Fig. 2.

For $\gamma_d = 18$, 3.2 and 1.4, respectively: the number of points in the filter function is 11, 17 and 8; χ_{ap}^2 is 1.29, 1.47 and 2.8; χ_{ag}^2 is 32, 111 and 19.5; and the lower limit on τ_{th} is 390 τ , 1,910 τ and 200 τ . These are the results we use in setting the lower limits. Performing the same analysis with the single component heating projections, the number of points in the filter function is 17, 19 and 9; χ_{ap}^2 is 1.5, 1.8 and 2.7; χ_{ag}^2 is 47, 125 and 21; and the lower limit on τ_{th} is 470 τ ,

1,690 τ and 220 τ . The differences in the lower limits obtained using three different data analyses (discussed in Supplementary Information) are in all cases less than 25%. We consider this to be the approximate systematic uncertainty in the lower limits.

Received 6 December 2005; accepted 28 February 2006.

- Coleman, S. Quantum sine-Gordon equation as the massive Thirring model. *Phys. Rev. D* **11**, 2088–2097 (1975).
- Tabor, M. *Chaos and Integrability in Nonlinear Dynamics* (Wiley, New York, 1989).
- Berman, G. P. & Izrailev, F. M. The Fermi–Pasta–Ulam problem: fifty years of progress. *Chaos* **15**, 015104 (2005).
- Berman, G. P., Borgonovi, F., Izrailev, F. M. & Smerzi, A. Irregular dynamics in a one-dimensional Bose system. *Phys. Rev. Lett.* **92**, 030404 (2004).
- Minguzzi, A. & Gangardt, D. M. Exact coherent states of a harmonically confined Tonks–Girardeau gas. *Phys. Rev. Lett.* **94**, 240404 (2005).
- Proukakis, N. P., Schmiedmayer, J. & Stoof, H. T. C. Quasi-condensate growth on an atom chip. Preprint at (<http://arxiv.org/cond-mat/0509154>) (2005).
- Herrmann, F. & Schmalzle, P. Simple explanation of a well-known collision experiment. *Am. J. Phys.* **49**, 761–764 (1981).
- Girardeau, M. Relationship between systems of impenetrable bosons and fermions in one dimension. *J. Math. Phys.* **1**, 516–523 (1960).
- Girardeau, M. D. & Wright, E. M. Measurement of one-particle correlations and momentum distributions for trapped 1D gases. *Phys. Rev. Lett.* **87**, 050403 (2001).
- Paredes, B. *et al.* Tonks–Girardeau gas of ultracold atoms in an optical lattice. *Nature* **429**, 277–281 (2004).
- Kinoshita, T., Wenger, T. & Weiss, D. S. Observation of a one-dimensional Tonks–Girardeau gas. *Science* **305**, 1125–1128 (2004).
- Rigol, M. & Muramatsu, A. Fermionization in an expanding 1D gas of hard-core bosons. *Phys. Rev. Lett.* **94**, 240403 (2005).
- Pedri, P., Santos, L., Öhberg, P. & Stringari, S. Violation of self-similarity in the expansion of a one-dimensional Bose gas. *Phys. Rev. A* **68**, 043601 (2003).
- Dunjko, V., Lorent, V. & Olshanii, M. Bosons in cigar-shaped traps: Thomas–Fermi regime, Tonks–Girardeau regime, and in between. *Phys. Rev. Lett.* **86**, 5413–5416 (2001).
- Kinoshita, T., Wenger, T. & Weiss, D. S. Local pair correlations in one-dimensional Bose gases. *Phys. Rev. Lett.* **95**, 190406 (2005).
- Gangardt, D. M. & Shlyapnikov, G. V. Stability and phase coherence of trapped 1D Bose gases. *Phys. Rev. Lett.* **90**, 010401 (2003).
- Tolra, B. L. *et al.* Observation of reduced three-body recombination in a correlated 1D degenerate Bose gas. *Phys. Rev. Lett.* **92**, 190401 (2004).
- Lieb, E. H. & Liniger, W. Exact analysis of an interacting Bose gas. The general solution and the ground state. *Phys. Rev.* **130**, 1605–1616 (1963).
- Wu, S., Wang, Y. J., Diot, Q. & Prentiss, M. Splitting matter waves using an optimized standing-wave light-pulse sequence. *Phys. Rev. A* **71**, 043602 (2005).
- Wang, Y. J. *et al.* Atom Michelson interferometer on a chip using a Bose–Einstein condensate. *Phys. Rev. Lett.* **94**, 090405 (2005).
- Moore, M. G., Bergeman, T. & Olshanii, M. Scattering in tight atom waveguides. *J. Phys. IV* **116**, 69–86 (2004).
- Olshanii, M. Atomic scattering in the presence of an external confinement and a gas of impenetrable bosons. *Phys. Rev. Lett.* **81**, 938–941 (1998).
- Wu, H. & Foot, C. J. Direct simulation of evaporative cooling. *J. Phys. B* **29**, L321–L328 (1996).
- Raman, C. *et al.* Evidence for a critical velocity in a Bose–Einstein condensed gas. *Phys. Rev. Lett.* **83**, 2502–2505 (1999).
- Carusotto, I. *et al.* Sensitive measurement of forces at the micron scale using Bloch oscillations of ultracold atoms. *Phys. Rev. Lett.* **95**, 093202 (2005).
- Fallani, L. *et al.* Observation of dynamical instability for a Bose–Einstein condensate in a moving 1D optical lattice. *Phys. Rev. Lett.* **93**, 140406 (2004).
- Gustavson, T. L., Landragin, A. & Kasevich, M. A. Rotation sensing with a dual atom-interferometer Sagnac gyroscope. *Class. Quantum Grav.* **17**, 2385–2398 (2000).
- Olshanii, M. & Dunjko, V. Interferometry in dense nonlinear media and interaction-induced loss of contrast in microfabricated atom interferometers. Preprint at (<http://arxiv.org/cond-mat/0505358>) (2005).
- Miesner, H.-J. *et al.* Bosonic stimulation in the formation of a Bose–Einstein condensate. *Science* **279**, 1005–1007 (1998).
- Kinoshita, T., Wenger, T. & Weiss, D. S. All-optical Bose–Einstein condensation using a compressible crossed dipole trap. *Phys. Rev. A* **71**, 11602(R) (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Olshanii, K. O'Hara, K. Gibble and J. Jain for discussions, and the NSF for support.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.S.W. (dsweiss@phys.psu.edu).

LETTERS

Parametric oscillation in vertical triple microcavities

C. Diederichs¹, J. Tignon¹, G. Dasbach¹, C. Ciuti¹, A. Lemaître², J. Bloch², Ph. Roussignol¹ & C. Delalande¹

Optical parametric oscillation is a nonlinear process that enables coherent generation of 'signal' and 'idler' waves, shifted in frequency from the pump wave^{1,2}. Efficient parametric conversion is the paradigm for the generation of twin or entangled photons for quantum optics applications such as quantum cryptography^{3,4}, or for the generation of new frequencies in spectral domains not accessible by existing devices. Rapid development in the field of quantum information requires monolithic, alignment-free sources that enable efficient coupling into optical fibres and possibly electrical injection. During the past decade, much effort has been devoted to the development of integrated devices for quantum information^{5–7} and to the realization of all-semiconductor parametric oscillators^{8,9}. Nevertheless, at present optical parametric oscillators typically rely on nonlinear crystals placed into complex external cavities, and pumped by powerful external lasers. Long interaction lengths are typically required and the phase mismatch between the parametric waves propagating at different velocities² results in poor parametric conversion efficiencies. Here we report the demonstration of parametric oscillation in a monolithic semiconductor triple microcavity with signal, pump and idler waves propagating along the vertical direction of the nanostructure. Alternatively, signal and idler beams can also be collected at finite angles, allowing the generation of entangled photon pairs. The pump threshold intensity is low enough to envisage the realization of an all-semiconductor electrically pumped micro-parametric oscillator.

In optical parametric oscillation, pump photons of frequency ω_p are converted into signal and idler photons at frequencies ω_s and ω_i (with $\omega_s + \omega_i = \omega_p$ for a second-order nonlinear interaction, or $\omega_s + \omega_i = 2\omega_p$ for a third-order interaction^{1,2}). It can be used, in particular, for generating highly correlated twin photons for quantum cryptography^{3,7}. Photon pairs can be used as pseudo-single-photon states^{5,6} in which one photon can trigger the detection of the second photon⁴. Alternatively, the two photons can be entangled like Einstein–Podolski–Rosen (EPR) states and can be used for quantum-key generation^{3,4}. Although they are highly desirable, the development of compact and easy to use quantum emitters based on parametric oscillation is at present limited by several factors: in common materials, owing to optical dispersion, the three waves propagate at different velocities. This gives rise to destructive interference and low parametric conversion efficiency.

Various strategies for retaining phase-matching for waves propagating over long interaction lengths have been proposed. The most common scheme relies on naturally birefringent materials. Other solutions were proposed¹⁰, such as random phase-matching¹¹ or quasi-phase-matching using structured systems with a periodically inverted nonlinearity^{9,12}. A common alternative relies on embedding the nonlinear material in an external cavity, resonant with the parametric frequencies¹³. Parametric oscillation can then be reached when the cavity losses become smaller than the parametric gain. Doubly or even triply resonant optical parametric oscillators (DROPO or TROPO) provide lower-threshold, efficient single-mode

operation and strongly correlated twin-beam sources¹⁴. Nevertheless, these systems are large and complex to operate.

Alternatively, monolithic semiconductor microcavities, in which the nonlinear material is clad between two highly reflecting Bragg mirrors, have recently been proposed to obtain cavity phase-matching for parametric oscillation in isotropic semiconductors with $\chi^{(2)}$ nonlinearity⁸. Moreover, it has been demonstrated that very large $\chi^{(3)}$ resonant polaritonic nonlinearities in semiconductor microcavities can be used to achieve low-threshold optical parametric oscillation^{15–17}. In this case, the strong coupling of excitons and photons—that is, polariton eigenstates¹⁸—is required. Here, phase-matching is achieved only thanks to the S-shaped energy–momentum polariton dispersion and for large pump angles^{15,19}. This has several important drawbacks. First, operation temperatures are intrinsically limited because the strong coupling and thus the phase-matching are lost for higher temperatures¹⁶, when the individual exciton and photon dispersions are recovered. Second, the large pump angle (close to the point of inflection of the polariton dispersion, typically $\theta_p = 15^\circ$) prevents the integration of a built-in electrical injection. Third, the idler wave is emitted at a large angle of around $2\theta_p$ and thus is only weakly coupled to light, precluding efficient collection.

Here, we demonstrate the realization of parametric oscillation in a monolithic semiconductor nanostructure that provides built-in phase-matching for $\theta_p = \theta_s = \theta_i = 0^\circ$. The signal and idler beams can also be collected at finite angles, meeting the requirements for the generation of entangled photon pairs. By engineering the photonic phase-space using coupled microcavities, triple resonance and phase-matching are simultaneously achieved, as suggested in ref. 20.

The scheme of the triply resonant microcavity nanostructure is shown in Fig. 1a: it is grown by molecular beam epitaxy and consists of three λ -GaAs microcavities ($\lambda = 840$ nm), coupled via two intermediate Bragg mirrors (distributed Bragg reflectors). The ensemble is enclosed between a bottom and a top Bragg mirror. The active medium consists of a set of quantum wells (QWs) grown in the cavities at the anti-nodes of the electromagnetic-field. We discuss here two samples showing distinct thresholds. In sample A, all mirrors are made of 13 pairs of quarter-wavelength layers of AlAs and GaAs. Each cavity contains a single 80 Å $\text{In}_{0.05}\text{Ga}_{0.95}\text{As}$ QW. The overall cavity finesse is 740. Sample B contains nine QWs per cavity and the mirrors are made of (from top to substrate) 15, 11, 11 and 18 pairs of quarter-wavelength layers, resulting in a larger finesse (1,350).

In Fig. 1b the transmission as a function of the incident angle (or equivalently the in-plane momentum k) is shown. It was calculated via standard transfer matrix formalism: in the left panel, a single cavity without QW absorption is considered, whereas in the right panel the three coupled cavities are resonant. Most importantly here, the finite transmission of the middle Bragg mirrors leads to an effective optical coupling that lifts the degeneracy of the cavity modes^{21–24}. These modes are delocalized throughout the whole structure. Hence, each QW is optically coupled to all three photonic

¹Laboratoire Pierre Aigrain, Ecole Normale Supérieure, 24 rue Lhomond, F-75231 Paris Cedex 05, France. ²LPN/CNRS, Route de Nozay, F-91460 Marcoussis, France.

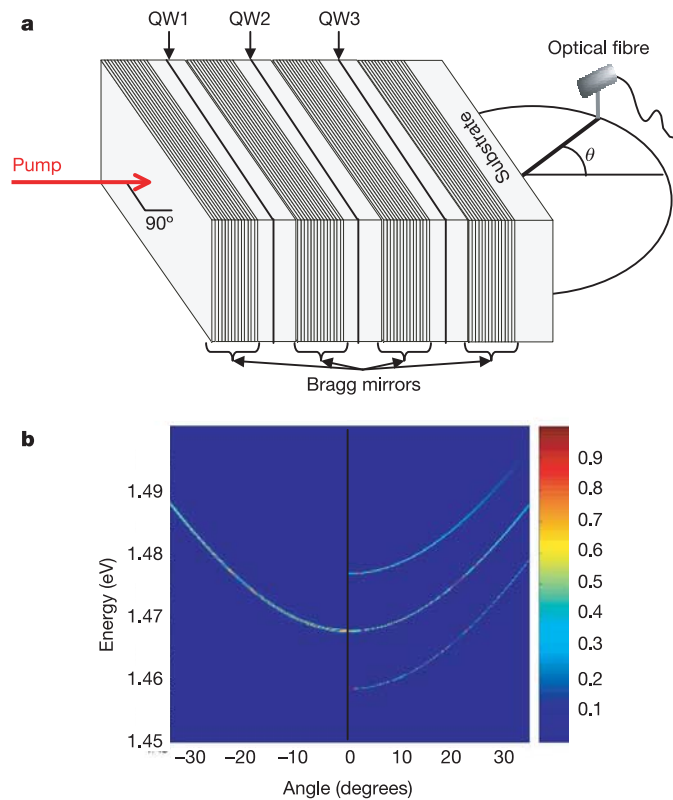


Figure 1 | Principle of the triple-microcavity structure. **a**, Sketch of a triple-microcavity structure (sample A) with three λ -GaAs-cavities optically coupled via two intermediate AlAs/GaAs distributed Bragg reflectors (13 pairs each). Each cavity contains a single 80 Å $\text{In}_{0.05}\text{Ga}_{0.95}\text{As}$ quantum well (QW). The substrate was polished to allow transmission. The pump-beam's incident angle is 0° with respect to the sample surface. A mechanical goniometer connected to an optical fibre allows us to measure the angle-resolved transmitted emission. **b**, Calculated transmission using the standard transfer matrix formalism showing the bare cavity mode (left) and the three coupled optical modes at resonance (right). The optical coupling lifts the degeneracy, resulting in optical normal modes delocalized throughout the whole structure. For clarity the QW absorption is set to zero.

modes. The exciton–photon detuning is defined as $\delta = E_\gamma - E_X$, where E_γ and E_X are the uncoupled cavity energy and the $1s$ exciton energy respectively (at 0°). A slight wedge has been voluntarily introduced in all three resonators (by interrupting the wafer rotation during the growth of the cavity spacers) so that the cavity-mode energies can be varied by moving the laser spot on the sample surface. Moreover, the wedge direction of the middle cavity is oriented perpendicularly to that of the outer cavities so that the cavity modes can be tuned relative to each other. The sample substrate has been polished to allow transmission experiments.

Figure 2 displays the measured angular-resolved emission at 6 K for samples A and B under continuous-wave resonant pumping of the middle energy mode (of energy $\hbar\omega_p$) at normal incidence. The position on each sample has been chosen so that all three cavities are at resonance and the resulting optical modes are split by the same energy. Above a threshold power, we observe the apparition of a strong, spectrally narrow, emission at a lower energy $\hbar\omega_s$ as well as higher energy $\hbar\omega_i$. The emission shows the characteristic signatures of parametric oscillation with energy ($2\hbar\omega_p = \hbar\omega_s + \hbar\omega_i$) and momentum ($2\mathbf{k}_p = \mathbf{k}_s + \mathbf{k}_i$) conservation. These conservation constraints nicely reflect in the (off-branch²⁵) negative dispersion of the idler mode (clearly visible in Fig. 2a for sample A), symmetric to the dispersion of the signal mode, in contrast to the positive dispersion of the conventional cavity mode. Given the low cavity finesse in this sample, the parametric scattering generates photon pairs of different

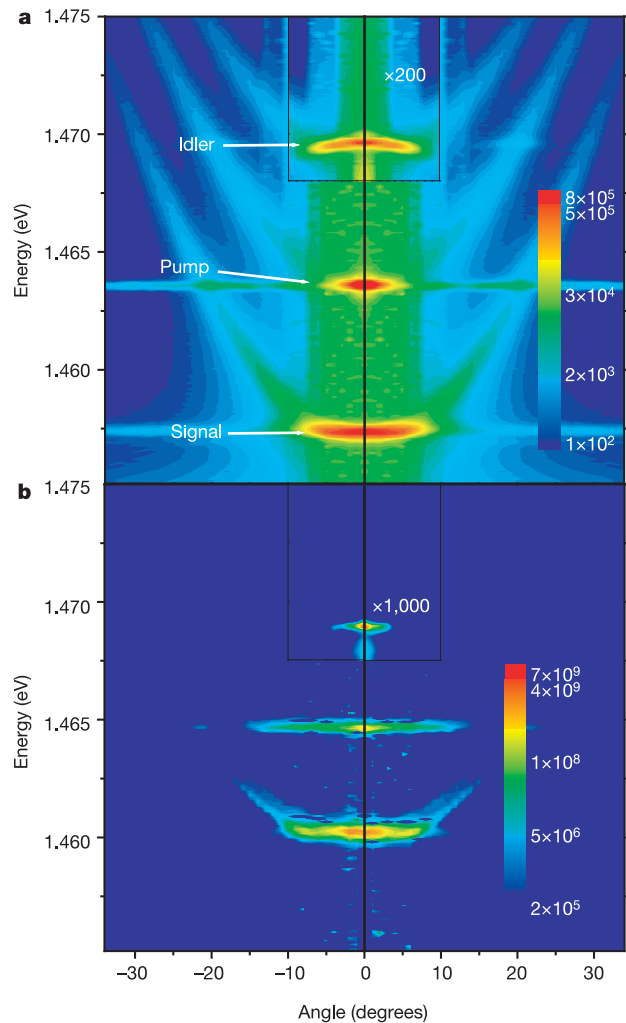


Figure 2 | Angle-resolved emission measured in transmission geometry. $T = 6$ K. The laser polarization is linear and a cross-polarized analyser allows us to minimize the laser transmission. **a**, Sample A (one QW per cavity, $E_x = 1,453.5$ meV, finesse 740). Parametric oscillation under resonant excitation of the second mode at $\hbar\omega_p = 1.4635$ eV and pump power 28 kW cm^{-2} . Detuning is $+10$ meV. Optical parametric oscillation is found above a threshold of 4.4 kW cm^{-2} with $\hbar\omega_s = 1.4574$ eV and $\hbar\omega_i = 1.4697$ eV. (The idler is multiplied by a factor of 200.) **b**, Same as **a** for sample B (nine QWs per cavity, $E_x = 1464.6$ meV, finesse 1,350). $\hbar\omega_p = 1.4646$ eV and pump power is 68 kW cm^{-2} . Detuning 0 meV. The threshold for sample B is 2.4 kW cm^{-2} . $\hbar\omega_s = 1.4602$ eV and $\hbar\omega_i = 1.4689$ eV. (The idler is multiplied by a factor of 1,000).

momentum within the mode linewidth. As a result, signal and idler photon pairs can also be collected at non-zero angles (that is, an idler photon at $\pm 7^\circ$ is correlated to a signal photon at $\mp 7^\circ$). This also matches the requirements for the emission of entangled photon pairs^{20,26}.

A threshold for optical parametric oscillator operation of 4.4 kW cm^{-2} is found for sample A. The threshold for sample B is lower (2.4 kW cm^{-2}), because sample B has a stronger gain medium and smaller cavity losses. The background emission clearly reveals a weak-coupling-like photon dispersion (see Fig. 2a) and not polariton dispersion. Thus in the excitation range used here, the strong coupling regime is bleached. In contrast to previously observed parametric oscillation in single microcavities^{15,16,19}, here, energy conservation and phase-matching is obtained thanks to the triple cavity resonance, without relying on the dispersion in the strong-coupling regime. Although it is not the focus of this Letter, we note that this parametric oscillation is also observed in the strong coupling

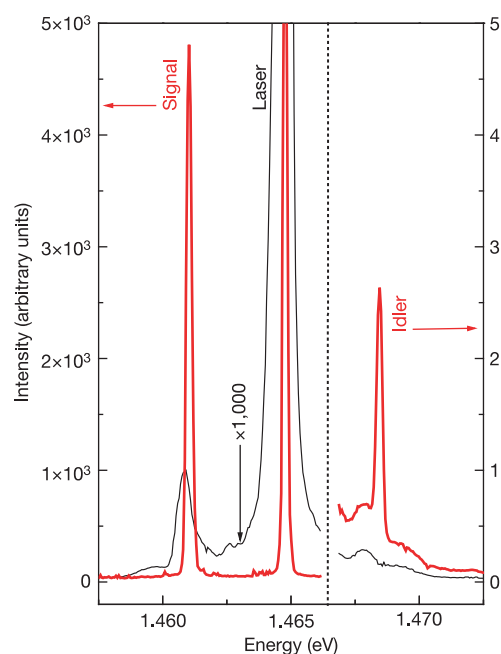


Figure 3 | Emission spectrum of the optical parametric oscillator. Emission spectrum at 0° below threshold (thin black line, power 2.4 kW cm^{-2}) and above the optical parametric oscillation threshold (bold red line), for a pump power of 3.2 kW cm^{-2} (sample B). The intensity is normalized with respect to the pump power so that the signal emission below threshold (multiplied by a factor of 1,000) is one.

regime (for low powers, specific detunings and temperatures below 50 K).

Figure 3 shows the emission spectrum measured at 0° on sample B, below threshold (black curve) and above threshold (red curve). The intensity is normalized with respect to the laser power: the signal intensity increases by a factor of 5,400 when the pump power is increased from 2.4 kW cm^{-2} to 3.2 kW cm^{-2} . Above threshold, the signal and idler emission narrow from 1 meV to below $200 \mu\text{eV}$ (resolution limited) and are blue-shifted (by 0.5 meV) with respect to the mode energy below threshold, as a consequence of the energy renormalization, as in the case of polariton parametric amplification¹⁹.

Last, Fig. 4 shows the signal and idler intensity versus pump power (sample B). The signal and idler emission shows an exponential behaviour above a threshold of 2.4 kW cm^{-2} followed by a quasi-linear power dependence. The signal to idler intensity ratio increases from two close to the threshold, up to three orders of magnitude for high powers when the idler saturates. The idler is quasi-degenerate with the QW absorption continuum, inducing losses and fast dephasing²⁶. Close to the threshold, this is well accounted for by the idler-to-signal linewidth ratio¹⁹. Nevertheless, this simple analysis does not apply for larger powers and the idler losses appear to be power-dependent. From a practical point of view for quantum optics applications, it is known that the attenuation of the idler will not reduce the visibility of the photon pair correlations but only the overall coincidence counting rate^{27,28}.

Figure 4 also shows (blue triangles) the emission of the low-energy mode when pumping resonantly the second mode for a position on the sample where energy conservation forbids parametric conversion (that is, strong deviation from triple resonance). The structure then behaves as a conventional VCSEL (vertical cavity surface emitting laser) with a larger threshold of 6 kW cm^{-2} (and no idler emission, nor energy conservation requirement). This again demonstrates the critical role of phase-matching in the up-conversion mechanism. It also shows that the parametric gain in this new type of structure can be made larger than the bare VCSEL gain.

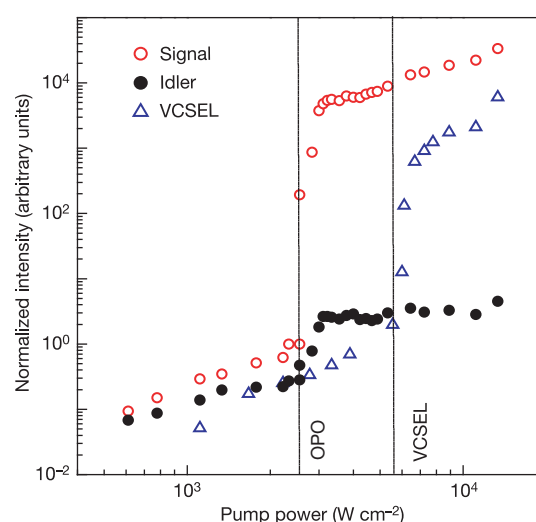


Figure 4 | Pump-power dependence of the optical parametric oscillator. Normalized (at the threshold) emitted intensity for the signal mode (red) and idler (black) for sample B. An optical parametric oscillation threshold is found at 2.4 kW cm^{-2} . Blue triangles indicate emission at the lowest-energy mode for an excitation of the middle-energy mode and for a position on the sample surface for which energy conservation does not allow parametric conversion. In this case, no idler emission is observed and a VCSEL-like laser emission at the signal energy is found above a threshold of 6 kW cm^{-2} (for the same exciton-photon detuning of 0 meV).

Further, parametric oscillation was observed above the temperature of liquid nitrogen, and is efficient at least up to 100 K (the samples were optimized for low-temperature operation and for larger temperatures, the QW energy is out of range). This clearly surpasses previous systems that rely on strong exciton-photon coupling, limited to 50 K (for InGaAs QWs and comparable materials¹⁶). Here, no intrinsic limitation remains for operation at room temperature and, as for VCSELs, proper design such as oxide confinement can be engineered for increased performances. Last, using 2-ps-long pulses with a 76-MHz repetition rate, we observed parametric oscillation in the same structures (not shown), with an ultralow threshold close to 0.5 nJ, typically three orders of magnitude lower than conventional optical parametric oscillators (see for example, ref. 9).

This micro-vertical TROPO has many similarities to VCSELs and has analogous advantages for optoelectronics applications. In particular, it has a high modulation bandwidth, the emission is single-mode, isotropic and well suited for coupling into optical fibres. Thanks to its ultralow threshold and ability to operate in continuous wave, it is easily compatible with electrical injection. This can be realized using three alternative approaches²⁹: (1) using an external injection VCSEL (resonant with the TROPO pump mode) attached to the optical parametric oscillator by wafer-bonding; (2) using a built-in VCSEL in the form of a fourth cavity, grown monolithically together with the optical parametric oscillator; (3) electrical injection of one of the three cavities is also possible (for example, the bottom one). This would require increasing the coupling Bragg mirror thickness for the pump cavity until the perturbative coupling to the other cavities is reached, so as not to perturb the signal and idler modes given by the two optically strongly coupled top cavities (see Supplementary Information for details).

Received 1 August 2005; accepted 19 January 2006.

1. Shen, Y. R. *The Principles of Nonlinear Optics* (Wiley Inter-Science, Hoboken, 2003).
2. Armstrong, J. A., Bloembergen, N., Ducuing, J. & Pershan, P. S. Interaction between light waves in a nonlinear dielectric. *Phys. Rev.* **127**, 1918–1939 (1962).

3. Jennewein, T., Simon, C., Weihs, G., Weinfurter, H. & Zeilinger, A. Quantum cryptography with entangled photons. *Phys. Rev. Lett.* **84**, 4729–4732 (2000).
4. Gisin, N., Robordy, G., Tittel, W. & Zbinden, H. Quantum cryptography. *Rev. Mod. Phys.* **74**, 145–195 (2002).
5. Santori, C., Fattal, D., Vuckovic, J., Solomon, G. S. & Yamamoto, Y. Indistinguishable photons from a single-photon device. *Nature* **419**, 594–597 (2002).
6. Moreau, E. *et al.* Single-mode solid-state single photon source based on isolated quantum dots in pillar microcavities. *Appl. Phys. Lett.* **79**, 2865–2867 (2001).
7. Tanzilli, S. *et al.* Highly efficient photon-pair source using a periodically poled lithium niobate waveguide. *Electron. Lett.* **37**, 26–28 (2001).
8. Haïdar, R., Forget, N. & Rosencher, E. Optical parametric oscillation in microcavities based on isotropic semiconductors: a theoretical study. *IEEE J. Quant. Elec.* **39**, 569–576 (2003).
9. Vodopyanov, K. L. *et al.* Optical parametric oscillation in quasi-phase-matched GaAs. *Opt. Lett.* **29**, 1912–1914 (2004).
10. Fiore, A., Berger, V., Rosencher, E., Bravetti, P. & Nagle, J. Phase matching using an isotropic nonlinear optical material. *Nature* **391**, 463–466 (1998).
11. Baudrier-Raybaut, M., Haïdar, R., Kupecek, Ph., Lemasson, Ph. & Rosencher, E. Random quasi-phase-matching in bulk polycrystalline isotropic nonlinear materials. *Nature* **432**, 374–376 (2004).
12. Myers, L. E., Miller, G. D., Eckardt, R. C., Fejer, M. M. & Byer, R. L. Quasi-phase-matched 1.064- μm -pumped optical parametric oscillator in bulk periodically poled LiNbO₃. *Opt. Lett.* **20**, 52–54 (1995).
13. Mertz, J., Debuisschert, T., Heidmann, A., Fabre, C. & Giacobino, E. Improvements in the observed intensity correlation of optical parametric oscillator twin beams. *Opt. Lett.* **16**, 1234–1236 (1991).
14. Teja, J. & Wong, N. C. Twin-beam generation in a triply resonant dual-cavity optical parametric oscillator. *Opt. Exp.* **2**, 65–71 (1998).
15. Savvidis, P. G. *et al.* Angle-resonant polariton amplifier. *Phys. Rev. Lett.* **84**, 1547–1550 (2000).
16. Saba, M. *et al.* High-temperature ultrafast polariton parametric amplification in semiconductor microcavities. *Nature* **414**, 731–735 (2001).
17. Dasbach, G. *et al.* Polarization inversion via parametric parametric scattering in quasi-one-dimensional microcavities. *Phys. Rev. B* **71**, 161308(R) (2005).
18. Weisbuch, C., Nishioka, M., Ishikawa, A. & Arakawa, Y. Observation of the coupled exciton-photon mode splitting in a semiconductor quantum microcavity. *Phys. Rev. Lett.* **69**, 3314–3317 (1992).
19. Ciuti, C., Schwendimann, P., Deveaud, B. & Quattropani, A. Theory of the angle-resonant polariton amplifier. *Phys. Rev. B* **62**, R4825–R4828 (2000).
20. Diederichs, C. & Tignon, J. Design for a triply resonant vertical emitting micro-optical parametric oscillator. *Appl. Phys. Lett.* **87**, 251107 (2005).
21. Stanley, R. P., Houdré, R., Oesterle, U., Ilegems, M. & Weisbuch, C. Coupled semiconductor microcavities. *Appl. Phys. Lett.* **65**, 2093–2095 (1994).
22. Pellandini, P. *et al.* Dual-wavelength laser emission from a coupled semiconductor microcavity. *Appl. Phys. Lett.* **71**, 864–866 (1997).
23. Michler, P., Hilpert, M. & Reiner, G. Dynamics of dual-wavelength emission from a coupled semiconductor microcavity laser. *Appl. Phys. Lett.* **70**, 2073–2075 (1997).
24. Armitage, A. *et al.* Optically induced splitting of bright excitonic states in coupled quantum microcavities. *Phys. Rev. B* **57**, 14877–14881 (1998).
25. Savvidis, P. G. *et al.* Off-branch polaritons and multiple scattering in semiconductor microcavities. *Phys. Rev. B* **64**, 075311 (2001).
26. Ciuti, C. Branch-entangled polariton pairs in planar microcavities and photonic wires. *Phys. Rev. B* **69**, 245304 (2004).
27. Ou, Z. Y. & Mandel, L. Further evidence of nonclassical behavior in optical interference. *Phys. Rev. Lett.* **62**, 2941–2944 (1989).
28. Ou, Z. Y. Quantum theory of fourth-order interference. *Phys. Rev. A* **37**, 1607–1619 (1988).
29. Tignon, J., Ciuti, C., Dasbach, G., & Diederichs, C. Micro-optical parametric oscillator with coupled cavities. French patent application number 05/05 708 (filed 06 June 2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was financially supported by the 'Clermont 2' project and the Délégation Générale pour l'Armement (DGA). LPA de l'ENS is 'Unité Mixte de Recherche Associée au CNRS et aux Universités Paris 6 et 7'.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper on www.nature.com/nature. Correspondence and requests for materials should be addressed to J.T. (Jerome.Tignon@lpa.ens.fr).

LETTERS

Management of singlet and triplet excitons for efficient white organic light-emitting devices

Yiru Sun¹, Noel C. Giebink¹, Hiroshi Kanno¹, Biwu Ma², Mark E. Thompson² & Stephen R. Forrest^{1†}

Lighting accounts for approximately 22 per cent of the electricity consumed in buildings in the United States, with 40 per cent of that amount consumed by inefficient ($\sim 15 \text{ lm W}^{-1}$) incandescent lamps^{1,2}. This has generated increased interest in the use of white electroluminescent organic light-emitting devices, owing to their potential for significantly improved efficiency over incandescent sources combined with low-cost, high-throughput manufacturability. The most impressive characteristics of such devices reported to date have been achieved in all-phosphor-doped devices, which have the potential for 100 per cent internal quantum efficiency²: the phosphorescent molecules harness the triplet excitons that constitute three-quarters of the bound electron-hole pairs that form during charge injection, and which (unlike the remaining singlet excitons) would otherwise recombine non-radiatively. Here we introduce a different device concept that exploits a blue fluorescent molecule in exchange for a phosphorescent dopant, in combination with green and red phosphor dopants, to yield high power efficiency and stable colour balance, while maintaining the potential for unity internal quantum efficiency. Two distinct modes of energy transfer within this device serve to channel nearly all of the triplet energy to the phosphorescent dopants, retaining the singlet energy exclusively on the blue fluorophore. Additionally, eliminating the exchange energy loss to the blue fluorophore allows for roughly 20 per cent increased power efficiency compared to a fully phosphorescent device. Our device challenges incandescent sources by exhibiting total external quantum and power efficiencies that peak at 18.7 ± 0.5 per cent and $37.6 \pm 0.6 \text{ lm W}^{-1}$, respectively, decreasing to 18.4 ± 0.5 per cent and $23.8 \pm 0.5 \text{ lm W}^{-1}$ at a high luminance of 500 cd m^{-2} .

Electrophosphorescent organic light-emitting devices (OLEDs) have been shown to harvest 100% of the excitons generated by electrical injection, corresponding to a fourfold increase in efficiency compared to that achievable in singlet-harvesting fluorescent OLEDs³. In this context, electrophosphorescent white OLEDs (WOLEDs) have been reported to exhibit^{4–7} high quantum (5–12%) and luminous power efficiencies ($6\text{--}20 \text{ lm W}^{-1}$) at brightnesses $< 100 \text{ cd m}^{-2}$. To date, however, blue electrophosphorescent devices have exhibited short operational lifetimes⁸ that limit the colour stability of all-phosphor-doped WOLEDs. Also, compared with their fluorescent counterparts, WOLEDs employing phosphorescent blue dopants excited via the conductive host introduce an approximately 0.8 eV exchange energy loss in power efficiency. This results from the energetic relaxation following intersystem crossing into the emissive triplet state. This loss can be avoided by resonant injection from the hole transport layer (HTL) and electron transport layer (ETL) into the phosphor triplet state^{6,9}, but the subsequent transfer to green and red dopants required to generate white light can

reintroduce these parasitic energy losses. Here we demonstrate a new WOLED architecture that uses a fluorescent emitting dopant to harness all electrically generated high energy singlet excitons for blue emission, and phosphorescent dopants to harvest the remainder of lower-energy triplet excitons for green and red emission.

This structure takes advantage of the fortuitous connection between the proportion of singlets dictated by spin statistics (that is, one singlet versus three triplets are produced by electrical excitation¹⁰) and the roughly 25% contribution of blue to the perceived white light spectrum. This allows for resonant energy transfer from both the host singlet and triplet energy levels that minimizes exchange energy losses, thereby maximizing device power efficiency while maintaining the potential for unity internal quantum efficiency (IQE). This approach has the further advantages of a stable white balance with current, a high efficiency at high brightness due to reduced geminate exciton recombination¹¹, and an enhanced lifetime due to the combined use of a stable fluorescent blue, and long lived phosphorescent green and red, dopants in a single emissive region.

The WOLED consists of a blue fluorophore, 4,4'-bis(9-ethyl-3-carbazovinylene)-1,1'-biphenyl (BCzVBi)¹², doped in a region spatially separate from the highly efficient green and red phosphorescent dopants fac-tris(2-phenylpyridine) iridium (Ir(ppy)_3) and iridium(III) bis(2-phenyl quinolyl-N,C^{2'}) acetylacetonate (PQIr), respectively. All lumophores are doped into a single, common conductive host, 4,4'-bis(N-carbazolyl)biphenyl (CBP), to form the extended emissive layer (EML) which is sandwiched between the electron transporting/hole blocking layer of 2,9-dimethyl-4,7-diphenyl-1,10-phenanthroline (BCP), and the 4,4'-bis[N-(1-naphthyl)-N-phenyl-amino]-biphenyl (α -NPD) HTL.

The principle of device operation is illustrated in Fig. 1. Excitons are formed on the host with a singlet-to-triplet formation ratio of χ_s/χ_t . Singlet excitons are transferred following a resonant Förster process onto the lightly doped (5%) blue fluorophore as opposed to direct trap formation¹². The non-radiative host triplets, however, cannot efficiently transfer to the fluorophore by the Förster mechanism, or by Dexter transfer owing to the low doping concentration. On the other hand, triplets typically have long diffusion lengths¹⁰ ($\sim 100 \text{ nm}$), and hence can migrate into the centre of the EML where they transfer onto the phosphors. Resonant transfer of the host triplet onto the green phosphor avoids exchange energy losses at this stage, although there are some unavoidable losses in transferring into the lowest energy red phosphor. Finally, placing an undoped host spacer with a thickness larger than the Förster radius ($\sim 3 \text{ nm}$) between the blue fluorophore and the phosphors prevents direct energy transfer from the blue dopant to the green and red phosphors. This device architecture is unique in that the singlet and triplet excitons are harvested along completely independent channels, and hence the transfer from host to dopant for both species can be separately

¹Department of Electrical Engineering, Princeton Institute for the Science and Technology of Materials (PRISM), Princeton University, Princeton, New Jersey 08544, USA.

²Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA. †Present address: Department of Electrical Engineering and Computer Science, Department of Physics, and Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA.

optimized to be nearly resonant, thereby minimizing energy losses while maintaining a unity IQE.

Figure 2 provides evidence for this transfer mechanism by comparing the un-normalized electroluminescent spectra of three devices at a current density of $J = 100 \text{ mA cm}^{-2}$. Device I has a 16-nm-thick CBP spacer placed between the two 5-nm-thick 5 wt% BCzVBi:CBP layers at each side of the EML, whereas device II has a 25-nm-thick uniformly doped 5 wt% BCzVBi:CBP EML. Both devices have nearly identical emission spectra and external quantum efficiencies (EQE) of $\eta_{\text{ext}} = (2.6 \pm 0.2)\%$, indicating that ostensibly 100% of the exciton formation occurs at the edges of the EML. Furthermore, lack of short wavelength CBP emission in device I suggests that the charge density in the middle region of EML available for exciton formation directly on the host is negligible. When a 20-nm-thick 3 wt% Ir(ppy)₃:CBP layer is inserted in the EML and separated from the two 5 wt% BCzVBi:CBP regions by 4-nm-thick undoped CBP spacers (device III), the total efficiency is increased to $\eta_{\text{ext}} = (5.2 \pm 0.2)\%$, with the additional 2.6% emission coming from Ir(ppy)₃. From this we conclude that exciton diffusion from the point of origin at the edges of the EML, rather than direct charge trapping and exciton formation on the phosphor, dominates, because carriers trapped by Ir(ppy)₃ would result in a noticeable decrease in the BCzVBi emission. Triplet diffusion from the edges of the EML to the phosphorescent doped region is consistent with previous observations in red fluorescent/phosphorescent OLEDs, where the fluorophore 'filters' out the singlet excitons, leaving only triplets to diffuse to a spatially separated phosphor doped region³.

To determine whether the location of the exciton formation region is predominantly at either the HTL/EML or EML/ETL interface, we compared the emission from two comparable devices with opposite symmetries, where the structure consisted of either NPD (30 nm)/5 wt% BCzVBi:CBP (10 nm)/CBP (20 nm)/BCP (40 nm), or NPD (30 nm)/CBP (20 nm)/5 wt% BCzVBi:CBP (10 nm)/BCP (40 nm). The corresponding maximum efficiencies of these devices

were $\eta_{\text{ext}} = (1.4 \pm 0.1)\%$ and $(1.8 \pm 0.1)\%$, both smaller than $\eta_{\text{ext}} = (2.6 \pm 0.2)\%$ for device II. Moreover, CBP emission at a peak wavelength of $\lambda = 390 \text{ nm}$ is observed in the first device, and α -NPD emission at $\lambda = 430 \text{ nm}$ is observed in the second structure. These observations suggest that excitons are generated at both the HTL/EML and EML/ETL interfaces, consistent with the ambipolar conductivity of CBP^{13,14}. Exciton formation at the edges, with correspondingly low generation in the bulk of the EML, can be understood as follows: large densities of holes (p) and electrons (n) pile up at the energy barriers at two EML interfaces. The exciton formation probability, which is $\sim n \times p$, is thus also significantly higher at these locations as compared with the EML bulk.

On the basis of these results, WOLEDs were fabricated by doping the middle region of the EML with both the green and red phosphorescent dopants (Fig. 3a, inset). As the high energy of the highest occupied molecular orbital of PQIr suggests that it can trap holes in the CBP host, a slight decrease of the blue emission intensity is observed in the WOLED. Fitting of the WOLED spectrum in Fig. 3b with the individual dopant spectra suggests that the ratio of emission from fluorescent to phosphorescent dopants approaches the ratio of 1/3, consistent with the singlet-to-triplet exciton formation ratio in emissive organic materials^{10,15,16}. Furthermore, given the performance characteristics of the purely fluorescent BCzVBi device (device II in Fig. 2), we also find that the fraction of excitons trapped by, and formed directly on, the phosphorescent dopants in the EML is $\chi_{\text{trap}} = (18 \pm 5)\%$ (see Methods). That is, approximately 20% of the excitons are formed by direct trapping on the phosphor dopants, whereas the remaining 80% are formed at the edges of the EML, at which point the triplets subsequently diffuse into the centre where they are transferred from host to phosphor dopant before emission.

A maximum forward viewing EQE of $\eta_{\text{ext}} = (11.0 \pm 0.3)\%$ is achieved at a current density $J = (1.0 \pm 0.6) \text{ mA cm}^{-2}$, and decreases only slightly to $\eta_{\text{ext}} = (10.8 \pm 0.3)\%$ at a high forward viewing luminance of 500 cd m^{-2} (Fig. 3a). This device gives a maximum forward viewing power efficiency of $\eta_p = (22.1 \pm 0.3) \text{ lm W}^{-1}$. As illumination sources are generally characterized by their total

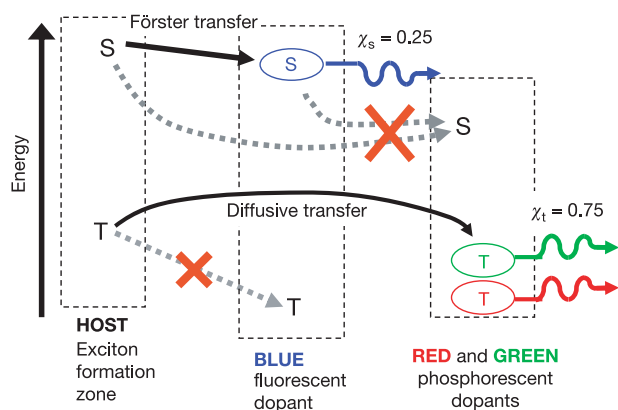


Figure 1 | Proposed energy transfer mechanisms in the fluorescent/phosphorescent WOLED. This illustrates the separate channels for triplet (T) and singlet (S) formation and transfer directly onto their corresponding emissive dopants. The majority of excitons are formed in the host material with a singlet-to-triplet formation ratio of χ_s/χ_t . The singlet excitons in the two formation regions at each side of the light emitting layer (EML) are rapidly, and near-resonantly, transferred to the blue fluorescent dye located in these regions. The phosphor-doped region is located in the centre of the EML and separated from the exciton formation zones by spacers of undoped host material. The triplets then diffuse efficiently to the central region, where they transfer to the lower energy green or red phosphor dopants, again by a nearly resonant process to the green dopant triplet manifold, and with some energy loss to the red triplet. Diffusion of singlet excitons to the phosphor dopants is negligible due to their intrinsically short diffusion lengths²³. Parasitic effects of charge trapping onto the phosphorescent dopants are discussed in the text.

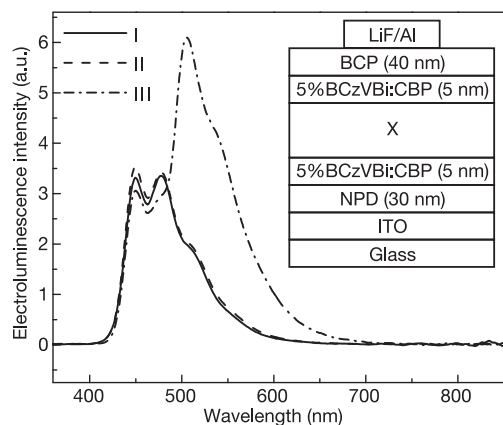


Figure 2 | Un-normalized electroluminescence spectra of three device structures shown in the inset. The spectra were measured at a current density of 100 mA cm^{-2} . Inset, schematic cross-section of the device; see text for definitions of abbreviations used. X = CBP (16 nm) for device I; X = 5 wt% BCzVBi:CBP (15 nm) for device II; X = CBP (4 nm)/3 wt% Ir(ppy)₃:CBP (20 nm)/CBP (4 nm) for device III. Quantitative comparison of the three spectra suggests that excitons are primarily formed at the two interfaces, and that fluorescent doping across the entire emission layer does not increase the blue luminescence intensity (compare devices I and II). However, doping the middle of the EML with the phosphor Ir(ppy)₃ results in additional green emission without a corresponding decrease in blue, BCzVBi emission (devices I and III), indicating that triplets formed on CBP are efficiently transferred to Ir(ppy)₃, whereas the BCzVBi acts to harvest, or 'filter out', all of the singlets.

emitted power, this device therefore has maximum total efficiencies⁶ of $\eta_{p,t} = (37.6 \pm 0.6) \text{ lm W}^{-1}$, and $\eta_{\text{ext},t} = (18.7 \pm 0.5)\%$. At a practical surface luminance of 500 cd m^{-2} , $\eta_{p,t} = (23.8 \pm 0.5) \text{ lm W}^{-1}$, or approximately 50% greater than for common incandescent lighting. Although the commercially available blue fluorophore has a low $\eta_{\text{ext}} = 2.7\%$ (compared with a maximum expected 5% achieved in the literature), the WOLED performance, nevertheless, represents a considerable improvement over the best all-phosphorescent devices previously reported^{6,7,17} (see Supplementary Information).

The intrinsic singlet-to-triplet ratio and the separation of the

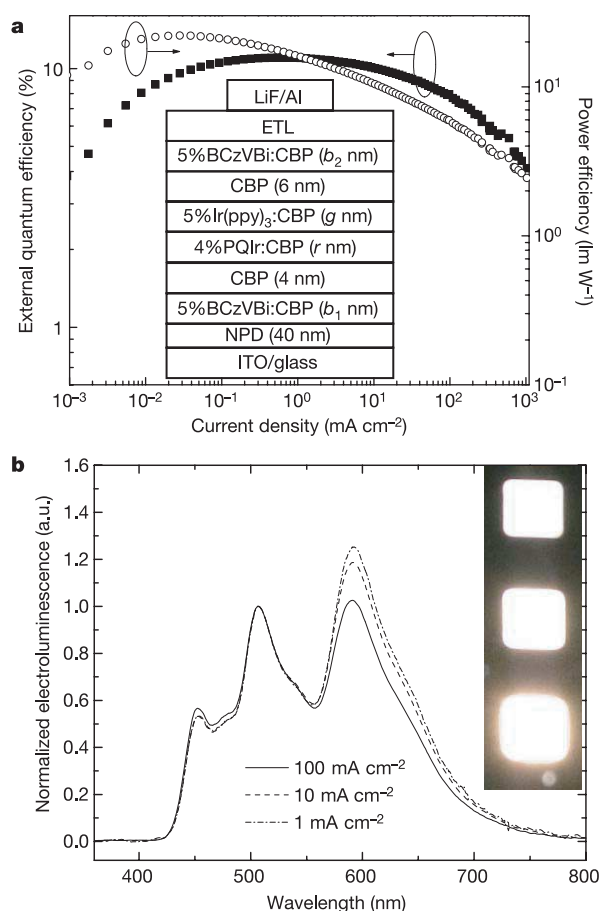


Figure 3 | Performance characteristics of the fluorescent/phosphorescent WOLED. **a**, Forward viewing external quantum efficiency (filled squares) and power efficiency (open circles) versus current density of the WOLED shown in the inset. The forward viewing external quantum efficiency peaks at $\eta_{\text{ext}} = (11.0 \pm 0.3)\%$ at $J = (1.0 \pm 0.6) \text{ mA cm}^{-2}$, and decreases slightly to $\eta_{\text{ext}} = (10.8 \pm 0.3)\%$ at a forward viewing luminance of 500 cd m^{-2} . The maximum forward viewing power efficiency is $\eta_p = (22.1 \pm 0.3) \text{ lm W}^{-1}$, with total peak and high luminance efficiencies of $\eta_{p,t} = (37.6 \pm 0.6)$ and $(23.8 \pm 0.5) \text{ lm W}^{-1}$ at 500 cd m^{-2} , respectively. The forward viewing luminance at 1 A cm^{-2} is $(83,000 \pm 7,000) \text{ cd m}^{-2}$. The drive voltage for this device is $(6.0 \pm 0.5) \text{ V}$ at $J = 10 \text{ mA cm}^{-2}$. Inset, schematic structure of the WOLED with the following layer thicknesses: $b_1 = 15 \text{ nm}$, $b_2 = 10 \text{ nm}$, $r = 8 \text{ nm}$, $g = 12 \text{ nm}$, and with an electron transport layer (ETL) consisting of 20-nm 4,7-diphenyl-1,10-phenanthroline (BPhen) followed by 20-nm Li doped BPhen in 1:1 molar ratio. Here, BPhen is used to further reduce device drive voltage. When 40-nm BCP is used as ETL and $b_1 = 10 \text{ nm}$, $b_2 = 10 \text{ nm}$, $r = 12 \text{ nm}$ and $g = 8 \text{ nm}$, η_{ext} and CRI are nearly identical to the above structure. **b**, Normalized electroluminescence spectra of WOLED emission at various current densities. Note that colour dependence on current density is minimal, with CRI = 85 at all three values of current density. Inset, images of three, 4.5 mm^2 devices, each driven at four times the drive current (from 1.7 to 28 mA cm^{-2}) of the device above it in the array (equivalent to a two f-stop difference in illumination) to show the colour stability of the emission.

channels in harvesting the two excitonic species gives a well-balanced and largely current-independent colour rendition, resulting in a colour rendering index of CRI = 85 at all current densities studied, which is the highest CRI among the reported values for a WOLED. The Commission Internationale d'Eclairage (CIE) coordinates have a negligible shift from (0.40, 0.41) at 1 mA cm^{-2} to (0.38, 0.40) at 100 mA cm^{-2} . This differs from observations of an all-phosphor-doped WOLED, where blue emission becomes stronger with increasing driving voltage⁶ owing to the requirement for high energy excitation of the blue phosphor. In the inset of Fig. 3b are images of three devices, each driven at 4 times higher drive current than the device above it in the array, to show the colour stability of the emission.

To further understand exciton diffusion, in Fig. 4a we plot (open circles) η_{ext} due to Ir(ppy)_3 emission versus the position (x) of a thin (5 nm) slab of 5 wt% Ir(ppy)_3 :CBP located at various distances from

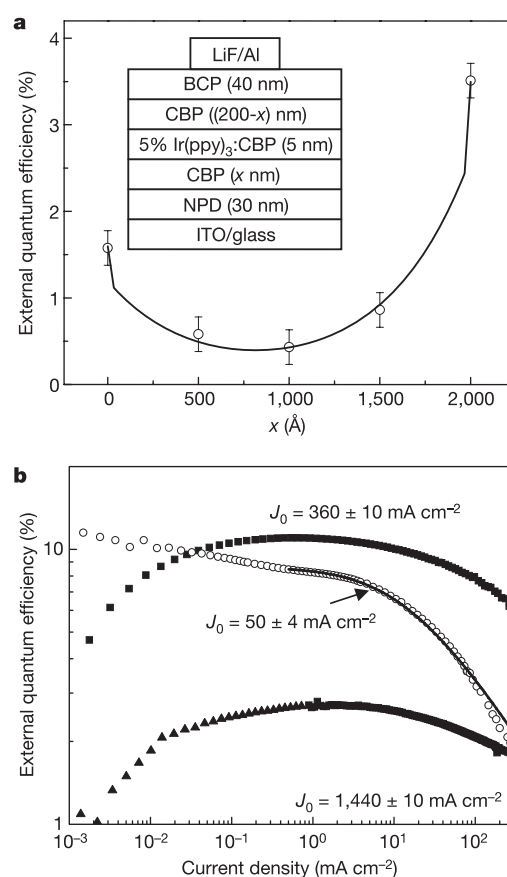


Figure 4 | Triplet diffusion profile and reduced efficiency roll-off at high currents. **a**, External quantum efficiency (open circles) from Ir(ppy)_3 emission at 10 mA cm^{-2} versus distance between the 50-Å slab of 5 wt% Ir(ppy)_3 :CBP and the NPD/CBP interface in the structure shown inset. A fit following equation (2) for triplet diffusion gives the solid curve and a triplet diffusion length of $L_D = (460 \pm 30) \text{ Å}$. The error bars indicate the standard deviations in measurement. Inset, schematic cross-section of the test structures: NPD (30 nm)/CBP ($x \text{ nm}$)/5 wt% Ir(ppy)_3 :CBP (5 nm)/CBP ((200 - x) nm)/BCP (40 nm), with $x = 0, 50, 100, 150, 200$. **b**, Comparison of external quantum efficiency roll-off. Open circles depict the performance of the all-phosphor white device of ref. 6, in comparison to the white device of this work (squares), and a blue fluorescent BCzVBi device (triangles). The high current roll-off of the phosphor device is described by triplet-triplet annihilation (fit shown as solid line), yielding an onset current density $J_0 = (50 \pm 4) \text{ mA cm}^{-2}$. The device of this work clearly demonstrates a roll-off that appears qualitatively similar to that of the all-fluorescent device. For comparison, $J_0 = (360 \pm 10) \text{ mA cm}^{-2}$ and $J_0 = (1,440 \pm 10) \text{ mA cm}^{-2}$ for the WOLED of this work and the all-fluorescent device, respectively.

the HTL/EML interface within a 200-nm-thick CBP EML (see Fig. 4a inset). Fitting (solid curve, see Methods) of the efficiency versus x yields a triplet diffusion length of $L_D = (460 \pm 30)$ Å, and predicts that $(75 \pm 5)\%$ of the phosphorescent emission results from triplet exciton diffusion from the adjacent EML interfaces, in agreement with the value calculated from analysis of the spectral content of the emission.

Compared with previous all-phosphor, high efficiency WOLEDs, the device also has a less pronounced efficiency roll-off at high current densities. For example, in Fig. 4b we show a comparison of η_{ext} versus J for an all-phosphor white device⁶, the device of this work, and a fluorescent BCzVBi device II. The high-current decline in η_{ext} of the all-phosphor white is due to triplet-triplet annihilation^{14,18}. In contrast, there is a striking resemblance between the efficiency roll-off of the current device, and that of device II. Modelling of the roll-offs in these two structures is complicated by recombination processes such as exciton-polaron quenching¹¹, singlet-triplet annihilation, and field-induced exciton dissociation. Nevertheless, for both of these latter devices, the current density at the point where η_{ext} has declined by half from its peak is >7 times that of the conventional phosphor device, while the peak EQE occurs at a value of J nearly 1,000 times larger. The apparent absence of triplet-triplet annihilation suggests that the highest density of triplet excitons is at the interfaces in the fluorescent doped regions, where they subsequently diffuse towards the centre, thereby lowering the local density (Fig. 4a) in the region of the guest phosphors. The reduced sensitivity of η_{ext} to current density is another clear difference between the WOLED of this study and previous, high efficiency all-phosphor devices.

We note that the efficiency of the present device can be further improved by using fluorescent dopants¹⁹ with IQE = 25% and phosphors^{20,21} with IQE = 100%, resulting in a total WOLED internal quantum efficiency of 100%. Using such 'ideal' chromophores, whose spectra are the same as the current dopants used, an approximately 34% total EQE and 60 lm W⁻¹ power efficiency can in principle be achieved using this structure, corresponding to a four-fold increase over incandescent power efficiency, and even competing with high efficiency, high CRI fluorescent lighting sources. As noted above, the exchange energy difference between the host singlet and dopant triplet states can lead to a loss of luminance efficiency in all-phosphor doped WOLEDs. By applying this design concept to systems where the host singlet is resonant with the blue fluorophore singlet state, and the host triplet is resonant with the green phosphor triplet level, this structure could have a power efficiency improvement of $\sim 20\%$ compared to similarly ideal all-phosphor devices. The highly efficient WOLED structure reported here, with a colour rendition that is unusually independent of current density, has potential for use in the next generation of sources for solid-state indoor lighting.

METHODS

Device manufacture. Devices were grown on clean glass substrates pre-coated with a 150-nm-thick layer of indium tin oxide (ITO) with a sheet resistance of 20 Ω per square. All organic layers were grown in succession without breaking vacuum ($\sim 10^{-7}$ torr). After organic film deposition, a shadow mask with 1-mm-diameter openings was affixed in a N₂ filled glove box before the cathode (consisting of 8-Å-thick LiF), followed by a 500-Å-thick Al cap, was deposited by high vacuum (10^{-6} torr) thermal evaporation. Current-voltage and EQE measurements were carried out using a semiconductor parameter analyser (HP 4145) and a calibrated Si photodiode (Hamamatsu S3584-08) following standard procedures²².

Data analysis. To interpret the emission spectrum, the WOLED EQE is expressed as:

$$\eta_{\text{ext}} = (1 - \chi_{\text{trap}})\eta_{\text{B}} + [(1 - \chi_{\text{trap}})\chi_{\text{t}} + \chi_{\text{trap}}]\eta_{\text{GR}} \quad (1)$$

where η_{B} and η_{GR} are respectively the EQEs of a singly doped blue fluorescent device and the comparable singly doped green and red phosphorescent devices, and χ_{trap} is the fraction of excitons trapped and formed directly on the phosphorescent dopants in the EML. By fitting the WOLED spectrum in

Fig. 3b at $J = 100 \text{ mA cm}^{-2}$ with the electroluminescence spectra of the three individual dopant materials, and accounting for photon energy in these power spectra, we find that $(20 \pm 2)\%$ of the total quantum efficiency is due to emission from the blue fluorescent dopant, and $(80 \pm 2)\%$ is from green and red phosphorescent dopants. Given the performance characteristics (η_{B}) of the purely BCzVBi device (device II in Fig. 2), we calculate $\chi_{\text{trap}} = (18 \pm 5)\%$ from the first term in equation (1).

Modelling. Exciton diffusion through the EML is modelled as shown in Fig. 4a (solid line) as follows: in steady state, and assuming that all singlet formation occurs at the HTL/EML ($x = 0$), and EML/ETL interfaces ($x = 200 \text{ nm}$) (in Fig. 4a), a solution to the triplet diffusion equation gives:

$$n(x) = \frac{1}{\sinh\left(\frac{L}{L_D}\right)} \left[\chi_{\text{t}} n_{\text{R}} \sinh\left(\frac{x}{L_D}\right) + \chi_{\text{t}} n_{\text{L}} \sinh\left(\frac{L-x}{L_D}\right) \right] + \chi_{\text{s}} n_{\text{L}} \delta(x) + \chi_{\text{s}} n_{\text{R}} \delta(x-L) \quad (2)$$

where n is the total exciton density: $n(x=0) = n_{\text{L}}$ and $n(x=L=200 \text{ nm}) = n_{\text{R}}$, L_D is the triplet diffusion length, and the delta function terms account for the presence of contributing singlets at the interfaces. As the EQE from Ir(ppy)₃ emission is proportional to the exciton density in the Ir(ppy)₃-doped slab, Fig. 4a (solid curve) shows the fit of the efficiency at $J = 10 \text{ mA cm}^{-2}$ versus x using equation (2) and $\chi_{\text{t}} = 3\chi_{\text{s}}$, from which we infer $L_D = (460 \pm 30)$ Å (error bars account for the spread in fits at additional current densities). With this calculated diffusion length, integration of the total exciton density in the phosphorescent doped region in the WOLED predicts that $(75 \pm 5)\%$ of the phosphorescent emission results from triplet exciton diffusion from the adjacent EML interfaces, in agreement with the value calculated from analysis of the spectral content of the emission (equation (1)).

Received 19 June 2005; accepted 13 February 2006.

- US Department of Energy *National Lighting Inventory and Energy Consumption Estimate* Vol. 1, xii (US Govt Printing Office, Washington DC, 2001).
- D'Andrade, B. W. & Forrest, S. R. White organic light-emitting devices for solid-state lighting. *Adv. Mater.* **16**, 1585–1595 (2004).
- Baldo, M. A. *et al.* Highly efficient phosphorescent emission from organic electroluminescent devices. *Nature* **395**, 151–154 (1998).
- Adamovich, V. *et al.* High efficiency single dopant white electrophosphorescent light emitting diodes. *N. J. Chem.* **26**, 1171–1178 (2002).
- D'Andrade, B. W. & Forrest, S. R. Effects of exciton and charge confinement on the performance of white organic p-i-n electrophosphorescent emissive excimer devices. *J. Appl. Phys.* **94**, 3101–3109 (2003).
- D'Andrade, B. W., Holmes, R. J. & Forrest, S. R. Efficient organic electrophosphorescent white-light-emitting device with a triple doped emissive layer. *Adv. Mater.* **16**, 624–627 (2004).
- Tokito, S., Iijima, T., Suzuki, T. & Sato, F. High-efficiency white phosphorescent organic light-emitting devices with greenish-blue and red-emitting layers. *Appl. Phys. Lett.* **83**, 2459–2461 (2003).
- Tung, Y. J. *et al.* A high efficiency phosphorescent white OLED for LCD backlight and display applications. *Proc. Soc. Inform. Display* **35**, 48–51 (2004).
- Holmes, R. J. *et al.* Efficient, deep-blue organic electrophosphorescence by guest charge trapping. *Appl. Phys. Lett.* **83**, 3818–3820 (2003).
- Baldo, M. A. Excitonic singlet-triplet ratio in a semiconducting organic thin film. *Phys. Rev. B* **66**, 14422–14428 (1999).
- Kalinowski, J. *et al.* Triplet energy exchange between fluorescent and phosphorescent organic molecules in a solid state matrix. *Chem. Phys.* **297**, 39–48 (2004).
- Hosokawa, C., Higashi, H., Nakamura, H. & Kusumoto, T. Highly efficient blue electroluminescence from a distyrylarylene emitting layer with a new dopant. *Appl. Phys. Lett.* **67**, 3853–3855 (1995).
- Adachi, C., Thompson, M. E. & Forrest, S. R. Architectures for efficient electrophosphorescent organic light-emitting devices. *IEEE J. Select. Top. Quant. Electron.* **8**, 372–377 (2002).
- Baldo, M. A. & Forrest, S. R. Transient analysis of organic electrophosphorescence: I. Transient analysis of triplet energy transfer. *Phys. Rev. B* **62**, 10958–10966 (2000).
- Reufer, M. *et al.* Spin-conserving carrier recombination in conjugated polymers. *Nature Mater.* **4**, 340–346 (2005).
- Segal, M., Baldo, M. A., Holmes, R. J., Forrest, S. R. & Soos, Z. G. Excitonic singlet-triplet ratios in molecular and polymeric organic materials. *Phys. Rev. B* **68**, 075211 (2003).
- Qin, D. S. & Tao, Y. White organic light-emitting diode comprising of blue fluorescence and red phosphorescence. *Appl. Phys. Lett.* **86**, 113507 (2005).
- Baldo, M. A., Adachi, C. & Forrest, S. R. Transient analysis of organic electrophosphorescence. II. Transient analysis of triplet-triplet annihilation. *Phys. Rev. B* **62**, 10967–10977 (2000).
- Murata, H., Kafafi, Z. H. & Uchida, M. Efficient organic light-emitting diodes

- with undoped active layers based on silole derivatives. *Appl. Phys. Lett.* **80**, 189–191 (2002).
20. Kawamura, Y. *et al.* 100% phosphorescence quantum efficiency of Ir(III) complexes in organic semiconductor films. *Appl. Phys. Lett.* **86**, 071104 (2005).
21. Adachi, C., Baldo, M. A., Thompson, M. E. & Forrest, S. R. Nearly 100% internal phosphorescence efficiency in an organic light emitting device. *J. Appl. Phys.* **90**, 5048–5051 (2001).
22. Forrest, S. R., Bradley, D. D. C. & Thompson, M. E. Measuring the efficiency of organic light-emitting devices. *Adv. Mater.* **15**, 1043–1048 (2003).
23. Sokolik, I., Priestley, R., Walser, A. D., Dorsinville, R. & Tang, C. W. Bimolecular reactions of singlet excitons in tris(8-hydroxyquinoline) aluminum. *Appl. Phys. Lett.* **69**, 4168–4170 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements H.K. is currently on leave from Sanyo Electric Co., Ltd., Osaka, Japan. The authors thank the Department of Energy and Universal Display Corp. for partial support of this work, as well as R. Holmes and B. W. D'Andrade for discussions.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.R.F. (forrest@princeton.edu).

Deep origin and hot melting of an Archaean orogenic peridotite massif in Norway

Dirk Spengler¹, Herman L. M. van Roermund¹, Martyn R. Drury¹, Luisa Ottolini², Paul R. D. Mason¹ & Gareth R. Davies³

The buoyancy and strength of sub-continental lithospheric mantle is thought to protect the oldest continental crust (cratons) from destruction by plate tectonic processes. The exact origin of the lithosphere below cratons is controversial, but seems clearly to be a residue remaining after the extraction of large amounts of melt^{1,2}. Models to explain highly melt-depleted but garnet-bearing rock compositions require multi-stage processes with garnet and clinopyroxene possibly of secondary origin^{1,3}. Here we report on orogenic peridotites (fragments of cratonic mantle incorporated into the crust during continent-continent plate collision⁴) from Otrøy, western Norway. We show that the peridotites underwent extensive melting during upwelling from depths of 350 kilometres or more, forming a garnet-bearing cratonic root in a single melting event. These peridotites appear to be the residue after Archaean aluminium depleted komatiite magmatism.

The discovery of microdiamonds in both orogenic peridotite and associated subducted crustal rocks in the Western Gneiss Region of Norway implies that both rock suites had a shared evolution during the Scandian orogeny 420–390 million years ago; when the Baltic continental crust was subducted into the diamond stability field (≥ 120 km)^{5–7}. However, the occurrence of pyroxene exsolution in subcontinental lithospheric mantle (SCLM) garnet after majorite indicates a deeper origin for the Western Gneiss Region peridotites (180–250 km)^{8,9} than for the surrounding crustal gneisses. Similar significant pressure differences occur in the Dabieshan-Sulu orogen, eastern China¹⁰ and the northern Tibetan plateau¹¹. Determining whether the majoritic ultrahigh-pressure microstructure originated before, or during, the Scandian orogeny has major consequences for the geodynamic interpretation^{6,8,12}. Here we demonstrate that the exsolution from majorite in the Otrøy peridotites is unrelated to the continental collision that brought the ultrahigh-pressure rocks to the surface.

The Ugelvik peridotite body on Otrøy Island is compositionally layered, comprising garnet-harzburgite and garnet-dunite (~65%), spinel-harzburgite and spinel-dunite (~33%) and minor amounts of garnet-pyroxenite and garnet-websterite layers (<2%)^{12,13}. Garnet-bearing layers contain typically 5–10% modal garnet, increasing locally up to 20%. Some layers of garnet-harzburgite and garnet-dunite contain isolated lenses of garnet-websterite¹⁴ (≤ 45 cm) and garnetite⁸ (≤ 25 cm in diameter). The term nodule is used for centimetre-scale polycrystalline garnet. Previously studied garnetite and garnet nodules^{8,9,15} contain pyroxene in two textural positions: grains of orthopyroxene (and rare olivine) in between garnet grains and needles of orthopyroxene (and minor clinopyroxene) in garnet cores, exemplified by sample DS0297 (Fig. 1a–d). Crystallographically oriented intracrystalline pyroxene has been accepted to be of exsolution origin from majoritic garnet¹⁶. Ugelvik garnetite DS0298

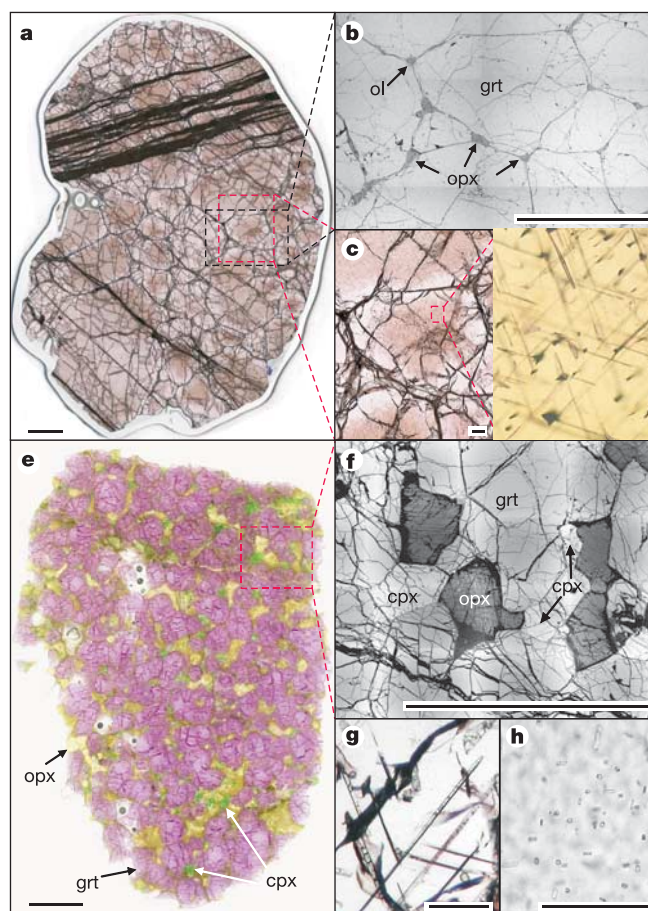


Figure 1 | Photomicrographs showing exsolved inter- and intracrystalline pyroxene within polycrystalline garnetites. a–d, Olivine-bearing garnetite DS0297 represents a well-described microstructure^{8,9,15} showing a mosaic garnet texture (a) with intercrystalline orthopyroxene and minor olivine (Mg# = 95.0) between individual garnet grains (b), crystallographically oriented, prismatic intracrystalline pyroxene needles in garnet cores (c, d) and needle-free garnet rims (e). e–h, Olivine-absent garnetite DS0298 differs with intercrystalline ortho- and clinopyroxene between individual garnet grains (e, f) and with intracrystalline pyroxene precipitates decreasing in size with the garnet grain size (g, h) in the garnet cores. a, e, Transmitted light; c, d, g, h, plane-polarized transmitted light; b, f, composite back-scattered electron image. Scale bars in a, b, e and f are 1 cm; in c, d, g and h they are 500 μ m. grt, garnet; opx, orthopyroxene; cpx, clinopyroxene; ol, olivine.

¹Faculteit Geowetenschappen, Universiteit Utrecht, Budapestlaan 4, 3584CD Utrecht, The Netherlands. ²Consiglio Nazionale delle Ricerche, Istituto di Geoscienze e Georisorse (IGG)-Sede di Pavia, Via Ferrata 1, 27100 Pavia, Italy. ³Faculteit der Aard- en Levenswetenschappen, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081HV Amsterdam, The Netherlands.

(Fig. 1e–h) differs in that it contains clino- and orthopyroxene, but no olivine, at intercrystalline positions. Intercrystalline pyroxene represents >20.6 vol% (Fig. 1e) with a orthopyroxene:clinopyroxene ratio of 4:1. If both inter- and intracrystalline pyroxenes were originally incorporated in the garnet structure, this rock must be derived from close to the Mantle Transition Zone (410–670 km).

To demonstrate that the garnetites were exhumed from this extraordinary depth we must establish that all pyroxene is cogenetic and not a result of an open system process such as melt addition. That all garnetite minerals are extremely depleted in light rare earth elements (REE) (Table 1 and Fig. 2b with chondrite-normalized $Ce/Sm < 0.08$) clearly demonstrates that no pyroxenes were in contact with a metasomatizing melt or fluid, nor were they crystallized directly from a melt—such processes would produce significantly less light-REE depletion¹⁷. The pyroxenes have inherited their light-REE-depleted character from the majoritic precursor. This conclusion is supported by the contrasting REE composition of clinopyroxene in a garnet-websterite lens (Fig. 2a), which occurs in the same peridotite body but is interpreted to be of melt origin^{6,14}.

A high-temperature exsolution origin of all clinopyroxene is suggested by intermineral REE partitioning. The equilibrium REE partitioning between garnet and clinopyroxene is strongly dependent upon atomic number, pressure and temperature^{18,19}, with a difference of about one order of magnitude in light REE between high and low temperatures¹⁸ (Fig. 2c). Analysed clinopyroxene in this work has Ce contents 8–13 times that of the host garnet. These values are at the high-temperature extreme of both high-temperature/high-pressure mineral partitioning experiments²⁰ and high-temperature ($\leq 1,380^\circ\text{C}$) peridotite xenoliths¹⁸ (Fig. 2a, c). The light-REE partition between garnet and clinopyroxene of both microstructures (inter- and intracrystalline, for the latter an exsolution origin seems

indisputable) indicates very high-temperature equilibration, $\geq 1,300^\circ\text{C}$.

Further support for an exsolution origin of intercrystalline pyroxene is provided by decompression experiments on garnet lherzolite performed up to 14 GPa that produced intercrystalline exsolved pyroxene and olivine from majoritic garnet¹⁵. We therefore conclude, on the basis of the chemistry and microstructure, that the intergrowth of clino- and orthopyroxene was exsolved at high-temperature ($\geq 1,300^\circ\text{C}$) from a majoritic garnet precursor. Recalculated precursor garnet compositions from the clinopyroxene-rich garnetite contains >19 mol% majorite (>3.19 silicon cations per formula unit at 12 oxygens), indicating that this garnet was stable at minimum (temperature-dependent) confining pressures of ≥ 11.5 GPa (refs 21, 22) corresponding to ≥ 350 km depth (Fig. 3).

Garnet nodules in harzburgite and dunite also constrain the melting history. Nodule DS0213 (~2 cm across) has garnet with millimetre-scale orthopyroxene inclusions. Garnet in nodule DS0289 (~5 cm) contains relicts of intracrystalline pyroxene needles, comparable to the microstructures in the larger garnetite samples. The two garnets are similar in colour and major element chemistry to garnet from the studied garnetites (Supplementary Table 1) except for higher Cr and lower Al. The nodular garnets are characterized by extremely depleted middle REE and exceptionally strongly fractionated heavy REE with chondrite-normalized $Dy/Yb < 0.07$ (Table 1, Fig. 2b). Garnet is the only significant REE-bearing phase in the garnet-harzburgites and garnet-dunites, and so garnet REE compositions mimic the whole rock.

Such a light-REE-depleted and fractionated-heavy-REE nature can only be explained by large degrees of melt extraction in the garnet stability field, $\geq 30\%$ (refs 23, 24). This conclusion is consistent with the high $Mg\#$ ($= Mg/(Mg + Fe)$) in olivine from both garnet-bearing and garnet-free harzburgites and dunites, with maxima of

Table 1 | Mineral REE data

Sample Mineral	DS0297 Core garnet†		DS0297 Intracrystalline clinopyroxene†		DS0298 Core garnet†		DS0298 Intracrystalline clinopyroxene†	
	Average (n = 3)	s.d.	Average (n = 9)	s.d.	Average (n = 5)	s.d.	Average (n = 4)	s.d.
La	0.001	0.001	0.052	0.014	0.002	0.002	0.017	0.006
Ce	0.009	0.002	0.15	0.05	0.005	0.004	0.061	0.027
Pr	NA	NA	NA	NA	NA	NA	NA	NA
Nd	0.09	0.03	0.40	0.07	0.030	0.004	0.25	0.11
Sm	0.22	0.05	0.50	0.03	0.07	0.02	0.28	0.13
Eu	0.15	0.02	0.33	0.05	0.09	0.02	0.18	0.04
Gd	0.80	0.08	0.62	0.04	0.65	0.04	0.33	0.08
Dy	2.57	0.35	0.40	0.03	2.26	0.13	0.44	0.06
Y	19.85	0.36	2.01	0.09	17.91	0.14	2.49	0.37
Er	3.05	0.25	0.18	0.02	2.55	0.13	0.20	0.01
Yb	4.39	0.22	0.17	0.03	3.30	0.09	0.27	0.01

Sample Mineral	DS0298 Core garnet‡		DS0298 Intercrystalline clinopyroxene‡		DS0213 Matrix garnet‡		DS0289 Matrix garnet‡	
	Average (n = 16)	s.d.	Average (n = 8)	s.d.	Average (n = 7)	s.d.	Average (n = 10)	s.d.
La	BD	BD	0.014	0.005	0.012	0.005	0.007	0.002
Ce	0.004	0.001	0.034	0.007	0.13	0.02	0.010	0.003
Pr	0.002	0.001	0.012	0.002	0.029	0.005	0.004	0.0005
Nd	0.05	0.01	0.14	0.01	0.11	0.03	0.046	0.012
Sm	0.09	0.02	0.26	0.02	0.031	0.009	0.033	0.008
Eu	0.08	0.02	0.13	0.01	0.011	0.004	0.010	0.002
Gd	0.59	0.07	0.43	0.04	0.023	0.004	0.035	0.007
Dy	2.23	0.14	0.45	0.02	0.059	0.014	0.091	0.009
Y	18.91	1.23	2.09	0.17	1.37	0.07	1.73	0.07
Er	2.43	0.19	0.18	0.02	0.23	0.02	0.29	0.03
Yb	3.45	0.22	0.11	0.03	0.75	0.09	0.89	0.05

Average REE data for garnet and intra- and inter-crystalline clinopyroxene with concentrations in p.p.m. and s.d. of 1 σ .

†Analysed by SIMS; ‡analysed by LA-ICP-MS. NA, not analysed; BD, below detection.

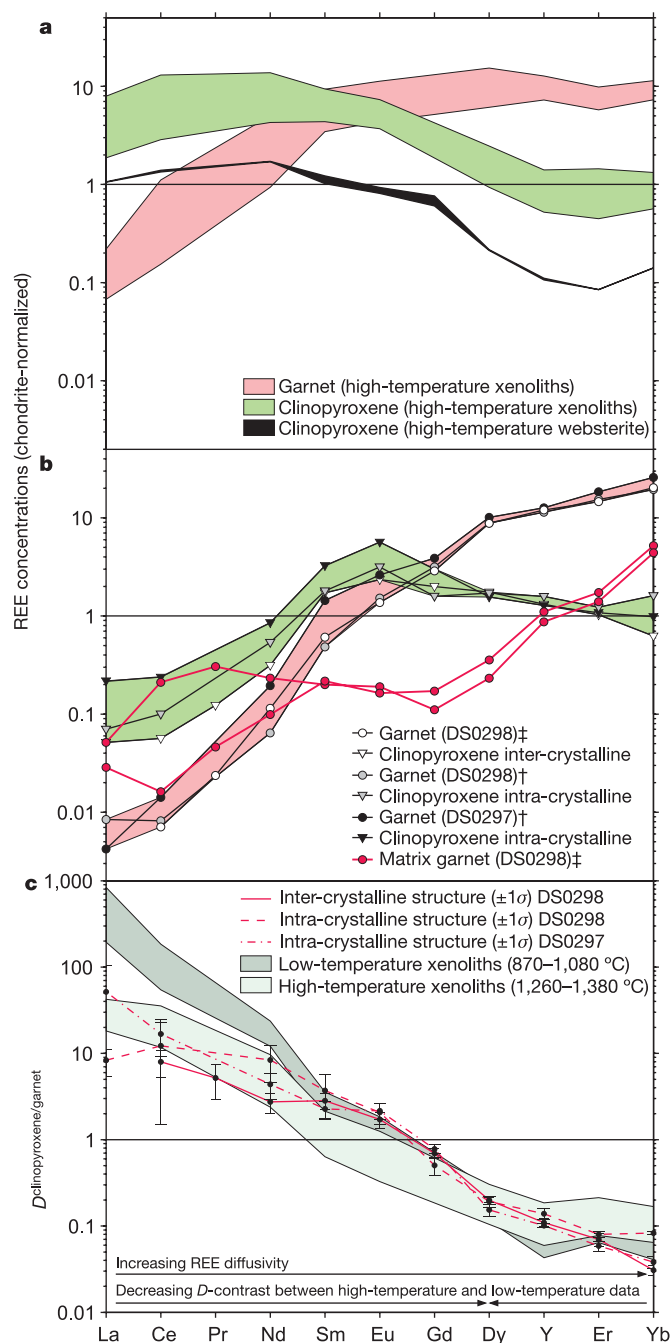


Figure 2 | C1-chondrite normalized mineral REE data from mantle xenoliths and Otrøy orogenic peridotites. **a**, Mineral REE concentrations from high-temperature xenoliths thought to be in chemical equilibrium¹⁸ and from a garnet-websterite lens on Otrøy, interpreted to be crystallized from a melt^{6,14}. Clinopyroxenes show elevated abundances of light REE. **b**, Mineral REE concentrations from Otrøy garnetites (shaded) and matrix garnets (red lines). All garnetite minerals show strongly depleted light REE (those analysed by SIMS marked with a dagger; those analysed by LA-ICP-MS marked with a double dagger) and contrast with nodular matrix garnet, which records even stronger depletion in middle REE and exceptionally fractionated heavy REE. **c**, Clinopyroxene–garnet REE partition coefficients D in Otrøy garnetites in comparison with those from high-temperature and low-temperature xenoliths¹⁸. Light REE in this study show a high-temperature equilibration, whereas middle and heavy REE indicate a low-temperature partitioning. The low-temperature REE partitioning is consistent with a re-equilibration during cooling below 1,100 °C over an appropriate time span because experimentally determined diffusion rates are faster for middle and heavy REE than for light REE¹⁹.

0.926–0.930 (refs 25, 32). Together, these data require the onset of melting during decompression at $\geq 1,800$ °C and ≥ 8 GPa (Fig. 3), otherwise garnet melts completely before olivine reaches these Mg#. Such high temperatures of melting only existed in the Archaean era^{25,26,32}.

Nd isotope ratios were determined to constrain the time of melting and cooling of the peridotites (Supplementary Table 2). Intercrystalline clinopyroxene and garnet in sample DS0298 define a mid-Proterozoic cooling age of 1.4 billion years (Gyr) before present (BP) (Supplementary Fig. 1). The 1.4 Gyr BP age establishes that intercrystalline pyroxene exsolution within the garnetite occurred long before the Scandian orogeny. Comparable REE partitioning between both intra- and intercrystalline pyroxene and garnet show that exsolution occurred at high temperature ($>1,300$ °C). We therefore disagree with previous suggestions that the Caledonian continental subduction was responsible for all decompression textures in the garnet peridotites⁶. Consequently we favour a multistage exhumation history^{8,27}. It follows that the occurrence of ultrahigh-pressure

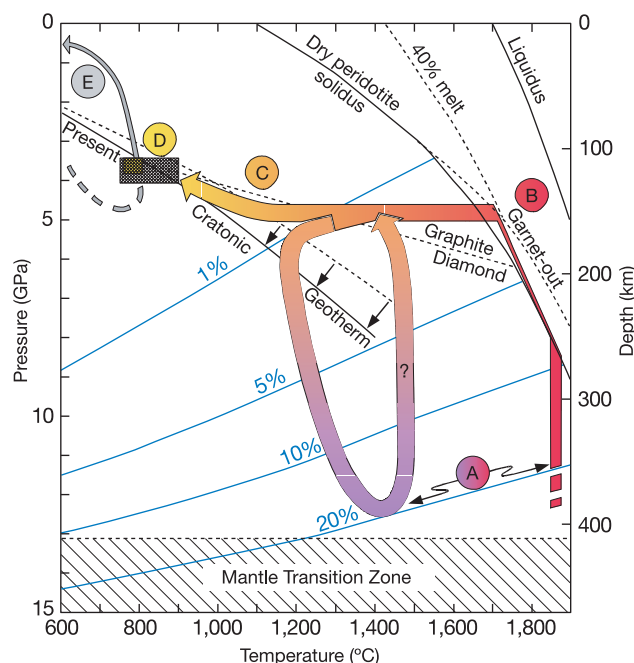


Figure 3 | Pressure-temperature diagram that schematically portrays the potential evolution of the Otrøy peridotites. Model 1 (red-orange-yellow path) requires near-isothermal decompression from depths at which the mantle contains $>19\%$ majorite in garnet (A). Extensive polybaric melting ($>30\%$) in the Archaean continued until the mantle with garnet in the residue reached the base of the overlying lithosphere (~ 150 km, B). Residua were stored and cooled towards the continental geotherm in the Proterozoic. Further decompression associated with the Svecofennian or Gothian orogeny cooled the peridotites further along the cooling continental geotherm, preserving the 1.4-Gyr-ago Nd mineral cooling age (C). Model 2 (red-orange-blue-yellow path) shows that after Archaean melting, residual peridotites are subducted to at least the Mantle Transition Zone and stored until thermal equilibration resulted in buoyancy-driven upwelling. Subsidiary upwelling took the peridotites to the base of the lithosphere (B–A–C) where cooling produced the 1.4-Gyr-ago Nd mineral cooling age (C). In both models further cooling until the Phanerozoic resulted in pressure-temperature conditions in equilibrium with the present continental geotherm (D). Final peridotite exhumation to the surface occurred during the Caledonian continent–continent collision along the grey path (E)^{4,31}. Rectangles represent equilibrated pressure-temperature conditions for garnet websterite lenses (grey cross-hatched) at Ugdevik (Otrøy) and Bardane (Fjortoft)^{6,7} and garnetites (yellow, see Supplementary Notes). Thin blue lines illustrate molar majorite components in solid solution with garnet²². The curve labelled ‘Garnet-out’ indicates the stability of garnet in peridotite below this curve³².

minerals in orogenic peridotites does not necessarily record conditions of deep continental subduction as recently argued^{16,10,11}.

The initial Nd isotope ratio of the mineral isochron, 0.5476 (Supplementary Table 2, Supplementary Fig. 1) is extremely radiogenic compared to mantle compositions at 1.4 Gyr BP ($\epsilon_{\text{Nd}}(1.4 \text{ Gyr BP}) = +719$). Nd model ages calculated relative to a chondritic evolution of the mantle are 2.9 Gyr BP and 2.5 Gyr BP for olivine-bearing and olivine-absent garnetite, respectively (Supplementary Table 2). These model ages indicate that the highly depleted garnet peridotites were formed in the Archaean era. Archaean melting of the garnet peridotites is also supported by a whole-rock Re–Os model age (3.3 Gyr BP) from Otrøy²⁸. The Otrøy peridotites therefore appear to be the first reported case of orogenic peridotites produced by extensive melt depletion entirely in the garnet stability field during the Archaean era.

There are two geodynamic models for SCLM formation that potentially explain the occurrence of peridotites derived from the Mantle Transition Zone at Otrøy. The simplest option (model 1, red-orange-yellow path in Fig. 3) is that Archaean adiabatic upwelling of mantle from the Mantle Transition Zone led to extensive melting and majorite exsolution. Melting stopped when the peridotites were underplated and accreted beneath an Archaean craton at depths of ~150 km. Over the next >1.0 Gyr the new SCLM cooled slowly but remained above the Nd closure temperature for clinopyroxene (~1,100 °C; ref. 19). Cooling below 1,100 °C occurred at 1.4 Gyr BP perhaps related to the Svecofennian or Gothian orogeny. The peridotites remained in the SCLM until exhumation 1.0 Gyr later during the continent–continent collision of the Caledonian orogeny.

A more complex alternative applies if the majorite formation postdates the melting (model 2, red-orange-blue-yellow path in Fig. 3). After hot upwelling, deep melting and accretion in the Archaean, part of the lower lithosphere was subducted, or delaminated, sinking to the Mantle Transition Zone accompanied by majorite formation. The low Fe peridotite residua subsequently underwent thermal equilibration and became buoyant²⁹, rising in a subsolidus upwelling to underplate existing SCLM in the Mid-Proterozoic, accompanied by majorite exsolution. This decompression could not involve partial melting, which would disrupt the Archaean Nd–Os isotope systematics.

Important implications of both models are that: (1) ultrahigh-pressure minerals in orogenic peridotites may be relicts that are not evidence for deep continental subduction and (2) deep asthenospheric underplating could form highly depleted SCLM with substantial amounts of garnet in the residue. We prefer the simplicity of model 1, but the more complex history of model 2 cannot be ruled out and indeed has been reproduced in numerical studies of early mantle convection³⁰. Model 2 implies that cratonic lithosphere retains the ancient isotopic signature of lithosphere that has been recycled back into the convecting mantle.

A fundamental characteristic of the Otrøy peridotites is that they preserve high Mg/Si ratios, low Ca/Al ratios (almost no clinopyroxene but 5–20% garnet) and relatively enriched heavy REE. These are all characteristics inferred for the source of Al-depleted komatiites²⁵.

METHODS

Major elements. Mineral major-element compositions were determined by electron microprobe analysis (EMPA) at Utrecht University (JEOL JXA8600, 15 kV, 20 nA, 30 s counting time in wavelength-dispersive spectrometry mode, correction for atomic number, X-ray absorption and fluorescence by using the ZAF routine, and external calibration against international standards).

Rare earth elements. Mineral REE concentrations were measured in garnet and clinopyroxene needles (sample DS0297 and DS0298) by secondary ion mass spectrometry (SIMS) using a Cameca IMS 4f ion microprobe at CNR-IGG (Pavia). Owing to the small size of clinopyroxene exsolution, we used a ¹⁶O[−] primary-ion current intensity of 1.5–3.5 nA. We used energy filtering (~100 V offset, 50 eV energy bandwidth). Data was acquired in five cycles using counting

times up to 125 s each for ¹³⁹La⁺, ¹⁴⁰Ce⁺, ¹⁴⁶Nd⁺ and ¹⁴⁹Sm⁺, 100 s each for ¹⁵¹Eu⁺, ¹⁵⁸Gd⁺, ¹⁶³Dy⁺, ¹⁶⁷Er⁺, ⁸⁹Y⁺ and ¹⁷⁴Yb⁺, and 15 s for ³⁰Si⁺. Calibration standards are the following: KAUG (Kakanui augite, New Zealand), KH1 (Kilbourne Hole clinopyroxene megacryst), MNAG (pyrope megacryst from Monastery Mine, South Africa) and KAKG (pyrope megacryst from Kakanui). Precision and accuracy of the method (tested on the standards) are ~10–20% at low p.p.m. concentrations. Below 1 p.p.m., precision is constrained by (Poisson) counting statistics and worsens to several tens of per cent. Under the adopted experimental conditions, La and Ce contents in garnet are comparable to detection limits for light REE.

Laser ablation inductively coupled mass spectrometry (LA-ICP-MS) was used at Utrecht University to measure mineral REE contents in garnet and clinopyroxene from the olivine-absent garnetite (sample DS0298) and in two porphyroclastic matrix garnets (from sample DS0213 and DS0289). The system comprises a GeoLas 200Q and a Micromass Platform ICP (193 nm excimer, beam diameter 120 µm, 10 Hz, 20 J cm^{−2}). Samples were ablated for 90–120 s per spot with signal integration over >60 s. Internal calibration was performed using Ca (EMPA) and external calibration was against the NIST SRM 612 glass standard. Gd was corrected for PrO interference.

Isotope geochemistry. Garnet and intercrystalline clinopyroxene were separated by hand picking, leached in warm HCl (1 M) and warm HNO₃ (1.4 M) each for 30 min, spiked and dissolved in hot concentrated HF and HNO₃. Sm and Nd were separated using standard ion exchange procedures. Sm and Nd abundances were determined by isotope dilution using a mixed 148/150 spike and analysed as metals by thermal ionisation mass spectrometry (TIMS) at the Vrije Universiteit Amsterdam using a Finnigan MAT 262. Nd isotope ratios on unspiked aliquots were measured as oxides. Replicate analysis of the La Jolla international standard yield 0.511863 ± 9 (*n* = 12).

Volume estimates. Photographs of back-scattered electron images and a high-resolution scanned thin section were contrast-enhanced using the two-dimensional analysis software 'analySIS' (Soft Imaging System; <http://www.soft-imaging.net/>). Threshold detections were applied to areas up to 2.1 mm² for intracrystalline pyroxene needles and up to 31 cm² for interstitial pyroxene (corresponding to the whole sample surface shown in Fig. 1e), and checked and corrected by hand. The resulting proportions of surface percentage of pyroxene and garnet were assumed to be equal to the proportion of volume percentage between the exsolved phases.

Received 1 November 2005; accepted 13 February 2006.

- Griffin, W. L. *et al.* Layered mantle lithosphere in the Lac de Gras area, Slave Craton: composition, structure and origin. *J. Petrol.* **40**, 705–727 (1999).
- Pearson, D. G., Canil, D. & Shirey, S. B. in *Treatise on Geochemistry*, Vol. 2: *The Core and Mantle* (eds Holland, H. D. & Turekian, K. K.) 171–275 (Elsevier, Oxford, 2003).
- Stachel, T., Viljoen, K. S., Brey, G. & Harris, J. W. Metasomatic processes in lherzolitic and harzburgitic domains of diamondiferous lithospheric mantle: REE in garnet from xenoliths and inclusions in diamonds. *Earth Planet. Sci. Lett.* **159**, 1–12 (1998).
- Brueckner, H. K. & Medaris, L. G. A general model for the intrusion and evolution of 'mantle' garnet peridotites in high-pressure and ultra-high-pressure terranes. *J. Metamorph. Geol.* **18**, 123–133 (2000).
- Dobrzhinetskaya, L. F. *et al.* Microdiamonds in high-grade metamorphic rocks from the Western Gneiss Region, Norway. *Geology* **23**, 597–600 (1995).
- Brueckner, H. K., Carswell, D. A. & Griffin, W. L. Paleozoic diamonds within a Precambrian peridotite lens in UHP gneisses of the Norwegian Caledonides. *Earth Planet. Sci. Lett.* **203**, 805–816 (2002).
- Van Roermund, H. L. M., Carswell, D. A., Drury, M. R. & Heijboer, T. C. Microdiamonds in a megacrystic garnet websterite pod from Bardane on the island of Fjortoft, western Norway: Evidence for diamond formation in mantle rocks during deep continental subduction. *Geology* **30**, 959–962 (2002).
- Van Roermund, H. L. M. & Drury, M. R. Ultra-high pressure (*P* > 6 GPa) garnet peridotites in western Norway: exhumation of mantle rocks from more than 185 km. *Terra Nova* **10**, 295–301 (1998).
- Van Roermund, H. L. M., Drury, M. R., Barnhoorn, A. & de Ronde, A. Relict majoritic garnet microstructures from ultra-deep orogenic peridotites in western Norway. *J. Petrol.* **42**, 117–130 (2001).
- Ye, K., Cong, B. & Ye, D. The possible subduction of continental material to depth greater than 200 km. *Nature* **407**, 734–736 (2000).
- Song, S., Zhang, L. & Niu, Y. Ultra-deep origin of garnet peridotite from the North Qaidam ultrahigh-pressure belt, Northern Tibetan Plateau, NW China. *Am. Mineral.* **89**, 1330–1336 (2004).
- Drury, M. R. *et al.* Emplacement of deep upper-mantle rocks into cratonic lithosphere by convection and diapiric upwelling. *J. Petrol.* **42**, 131–140 (2001).
- Carswell, D. A. Possible primary upper mantle peridotite in Norwegian basal gneiss. *Lithos* **1**, 322–355 (1968).

14. Carswell, D. A. Garnet pyroxenite lens within Ugelvik layered garnet peridotite. *Earth Planet. Sci. Lett.* **20**, 347–352 (1973).
15. Dobrzhinetskaya, L. F. *et al.* Precipitation of pyroxenes and olivine from majoritic garnet: simulation of peridotite exhumation from great depth. *Terra Nova* **16**, 325–330 (2004).
16. Haggerty, S. E. & Sautter, V. Ultradeep (greater than 300 kilometers), ultramafic upper mantle xenoliths. *Science* **248**, 993–996 (1990).
17. Burgess, S. R. & Harte, B. Tracing lithosphere evolution through the analysis of heterogeneous G9–G10 garnets in peridotite xenoliths, II: REE chemistry. *J. Petrol.* **45**, 609–634 (2004).
18. Schmidberger, S. S. & Francis, D. Constraints on the trace element composition of the Archean mantle root beneath Somerset Island, Arctic Canada. *J. Petrol.* **42**, 1095–1117 (2001).
19. Van Orman, J. R., Grove, T. L., Shimizu, N. & Layne, G. D. Rare earth element diffusion in a natural pyrope single crystal at 2.8 GPa. *Contrib. Mineral. Petrol.* **142**, 416–424 (2002).
20. Johnson, K. T. M. Experimental determination of partition coefficients for rare earth and high-field-strength elements between clinopyroxene, garnet, and basaltic melt at high pressures. *Contrib. Mineral. Petrol.* **133**, 60–68 (1998).
21. Fei, Y. & Bertka, C. M. in *Mantle Petrology: Field Observations and High-Pressure Experimentation* (eds Fei, Y., Bertka, C. M. & Mysen, B. O.) Spec. Publ. No. 6, 189–207 (The Geochemical Society, Houston, 1999).
22. Gasparik, T. *Phase Diagrams for Geoscientists—An Atlas of the Earth's Interior* (Springer, Berlin, 2003).
23. Hellebrand, E., Snow, J. E., Hoppe, P. & Hofmann, A. W. Garnet-field melting and late stage refertilization in 'residual' abyssal peridotites from the central indian ridge. *J. Petrol.* **43**, 2305–2338 (2002).
24. Ionov, D. Chemical variations in peridotite xenoliths from Vitim, Siberia: inferences for REE and Hf behaviour in the garnet-facies upper mantle. *J. Petrol.* **45**, 343–367 (2004).
25. Herzberg, C. Geodynamic information in peridotite petrology. *J. Petrol.* **45**, 2507–2530 (2004).
26. Takahashi, E. Speculations on the Archean mantle: missing link between komatiite and depleted garnet peridotite. *J. Geophys. Res.* **95**, 15941–15954 (1990).
27. Green, H. W. II, Dobrzhinetskaya, L. & Bozhilov, K. N. Mineralogical and experimental evidence for very deep exhumation from subduction zones. *J. Geodynam.* **30**, 61–76 (2000).
28. Beyer, E. E., Brueckner, H. K., Griffin, W. L., O'Reilly, S. Y. & Graham, S. Archean mantle fragments in Proterozoic crust, Western Gneiss Region, Norway. *Geology* **32**, 609–612 (2004).
29. Ringwood, A. E. Role of the transition zone and 660 km discontinuity in mantle dynamics. *Phys. Earth Planet. Inter.* **86**, 5–24 (1994).
30. De Smet, J. H., Van den Berg, A. P. & Vlaar, N. J. Early formation and long-term stability of continents resulting from decompression melting in a convecting mantle. *Tectonophysics* **322**, 19–33 (2000).
31. Terry, M. P., Robinson, P. & Krogh Ravna, E. J. Kyanite eclogite thermobarometry and evidence for thrusting of UHP over HP metamorphic rocks, Nordøyane, Western Gneiss Region, Norway. *Am. Mineral.* **85**, 1637–1650 (2000).
32. Walter, M. J. in *Treatise on Geochemistry, Vol. 2: The Core and Mantle* (eds Holland, H. D. & Turekian, K. K.) 363–394 (Elsevier, Oxford, 2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This project was initiated during the NWO-PIONIER project of M.R.D. and subsequently funded by the Utrecht University Institute of Geodynamic Research. Electron microscopy was conducted at the Electron Microscopy Utrecht centre. The National Research Council of Italy is acknowledged for funding the SIMS laboratory at CNR-IGG (Pavia). The isotope and LA-ICP-MS facilities at Vrije Universiteit and Universiteit Utrecht are supported by the Netherlands Organisation for Scientific Research. We thank M. Røberg and T. Wilhelmsen for providing appropriate samples. P. J. O'Brien is thanked for suggestions that have led to improvements of this paper. Thanks also to H. K. Brueckner, D. A. Carswell, H. Green II, A. Scherstén and W. van Westrenen.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.S. (spengler@geo.uu.nl).

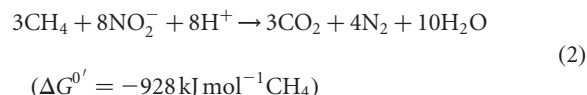
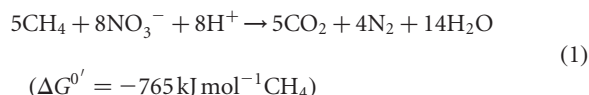
LETTERS

A microbial consortium couples anaerobic methane oxidation to denitrification

Ashna A. Raghoebarsing¹, Arjan Pol¹, Katinka T. van de Pas-Schoonen¹, Alfons J. P. Smolders², Katharina F. Ettwig¹, W. Irene C. Rijpstra³, Stefan Schouten³, Jaap S. Sinninghe Damsté³, Huub J. M. Op den Camp¹, Mike S. M. Jetten¹ & Marc Strous¹

Modern agriculture has accelerated biological methane and nitrogen cycling on a global scale^{1,2}. Freshwater sediments often receive increased downward fluxes of nitrate from agricultural runoff and upward fluxes of methane generated by anaerobic decomposition³. In theory, prokaryotes should be capable of using nitrate to oxidize methane anaerobically, but such organisms have neither been observed in nature nor isolated in the laboratory^{4–8}. Microbial oxidation of methane is thus believed to proceed only with oxygen or sulphate^{9,10}. Here we show that the direct, anaerobic oxidation of methane coupled to denitrification of nitrate is possible. A microbial consortium, enriched from anoxic sediments, oxidized methane to carbon dioxide coupled to denitrification in the complete absence of oxygen. This consortium consisted of two microorganisms, a bacterium representing a phylum without any cultured species and an archaeon distantly related to marine methanotrophic Archaea. The detection of relatives of these prokaryotes in different freshwater ecosystems worldwide^{11–14} indicates that the reaction presented here may make a substantial contribution to biological methane and nitrogen cycles.

Global biogeochemical cycles are mainly driven by microorganisms feeding on one-carbon compounds such as methane or carbon dioxide. Each step in the element cycles is catalysed by a specific group of microorganisms. These may or may not be evolutionarily related, but they share a similar lifestyle and so form an 'ecological guild'. Thermodynamic calculations show that most of these guilds have already been discovered (Supplementary Fig. S1), but the microorganisms that couple the anaerobic oxidation of methane (AOM) to denitrification, shown in equations (1) and (2), are considered missing in nature^{4–8}:



As AOM coupled to denitrification is possible in theory, both thermodynamically and biochemically (through reverse methanogenesis^{15,16}), such microorganisms might in fact exist and consequently our understanding of biogeochemical methane cycling may be incomplete. The lack of experimental evidence for the occurrence of AOM coupled to denitrification is perhaps not surprising, because this process would be expected to occur close to the oxic/anoxic interface in sediments. This interface is generally characterized by

steep gradients, occurring within millimetres, masking the process from geochemical detection. Furthermore, laboratory enrichment of the responsible microorganisms could be difficult because of their potentially very slow growth¹⁷.

Here we report the successful enrichment of consortia of microorganisms capable of AOM coupled to denitrification. Anoxic sediment from the Twentekanaal, a canal in the Netherlands, was used as the inoculum for the enrichment culture. This canal contained nitrate concentrations of up to 1 mM and the sediment was saturated with methane, which is typical for freshwater habitats receiving agricultural runoff. A one-litre sample from the upper layer of the sediment was incubated anoxically in the laboratory. Methane was supplied as the only electron donor, and mineral medium containing nitrate (NO_3^-), nitrite (NO_2^-), bicarbonate and trace elements was supplied and removed continuously. Over 16 months of incubation, the influent nitrite concentration was gradually increased to 6 mM, but the actual concentration in the culture medium remained at about 0.1 mM, indicating the growth of a microbial population consuming nitrite. Some nitrate ($<1 \text{ mM}$) was also consumed. Methane consumption could not yet be observed experimentally because it was supplied in large excess and any potential conversion remained within the error margin for the CH_4 analysis.

To measure methane consumption, the media and methane supply were stopped and excess methane was removed by flushing with helium gas. The consumption of methane, nitrite and nitrate was now apparent and dinitrogen gas evolved (Fig. 1). Nitrite and nitrate together accounted for all of the produced dinitrogen gas. However, the total denitrification rate ($28.8 \pm 2 \mu\text{mol N}_2 \text{ h}^{-1}$) was not completely accounted for by methane oxidation ($13.4 \pm 1 \mu\text{mol CH}_4 \text{ h}^{-1}$) according to equations (1) and (2). Control experiments indicated that this difference was caused by the oxidation of organic compounds from the inoculum or the mineral medium. In these control experiments, the denitrification rate in the absence of methane was $5.5 \pm 0.5 \mu\text{mol N}_2 \text{ h}^{-1}$, and on methane addition it increased to $21.5 \pm 2 \mu\text{mol N}_2 \text{ h}^{-1}$. As methane itself was consumed at $22.0 \pm 2 \mu\text{mol CH}_4 \text{ h}^{-1}$, the stoichiometry of AOM was almost completely consistent with the above equations.

The enrichment culture used nitrite in preference to nitrate as the substrate for denitrification—for experiments in which nitrite was depleted in the presence of both methane and nitrate, AOM ceased, and resumed only when nitrite was re-introduced 2 h after depletion. However, during longer incubation periods in the absence of nitrite (10–20 h), AOM restarted and was then coupled to nitrate consumption. This suggests that the enrichment culture could adapt to nitrate, and that both nitrite and nitrate were suitable substrates for AOM

¹Department of Microbiology, and ²Department of Aquatic Ecology and Environmental Biology, Institute for Water and Wetland Research, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands. ³Royal Netherlands Institute for Sea Research (NIOZ), Department of Marine Biogeochemistry and Toxicology, PO Box 59, 1790 AB Den Burg, The Netherlands.

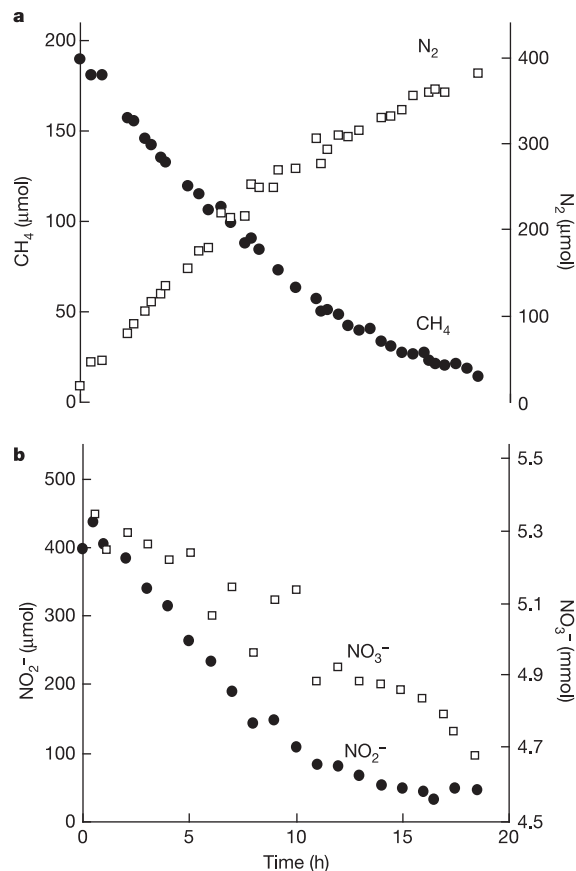


Figure 1 | AOM is coupled to the denitrification of nitrite by the enrichment culture after 16 months of enrichment. **a, b,** The total amounts of methane, dinitrogen gas, nitrate and nitrite present in the culture vessel are indicated. The initial concentrations of these compounds in the culture liquid were 6.0 μM, 0.30 μM, 3.6 mM and 0.24 mM, respectively. During this experiment, the total amount of protein in the enrichment culture was 100 mg.

(equations (1) and (2)). The apparent affinity constant for methane was very high (less than 0.6 μM, Fig. 1). The affinity of sulphate-dependent AOM for methane is four orders of magnitude lower¹⁸ (affinity constant >16 mM). The addition of 2 mM sulphate to the

enrichment culture neither stimulated nor inhibited AOM. Together, these experiments show unambiguously that methane can be oxidized anaerobically in this system, and that this oxidation is coupled to denitrification. The participation of oxygen in this process can be fully excluded. First, we measured oxygen levels continuously in the culture liquid and periodically in the headspace, but did not detect any (detection limit 80 p.p.m.). Second, all of the detected dinitrogen gas in the headspace of the culture was accounted for by the consumption of nitrite and nitrate. Had air leaked into the culture, more dinitrogen gas would have been detected in the headspace than was predicted by the reaction stoichiometry. Third, the stoichiometry of methane consumption coupled to denitrification was in good agreement with equations (1) and (2). Had oxygen been involved, much less nitrite or nitrate would have been consumed per mol of methane. Finally, no methane was consumed as the enrichment culture adapted from nitrite to nitrate (see above).

We investigated the incorporation of methane into microbial biomass by analysing the membrane lipids of the enrichment culture. We detected a single archaeal and several bacterial biomarkers (Table 1). The archaeal biomarker sn2-hydroxyarchaeol, found in methanogens¹⁹ and ANME-2 methanotrophic Archaea²⁰, was the only biomarker substantially depleted in ¹³C compared to methane. This change indicates that the carbon in this marker compound originated from methane. When ¹³C-labelled methane was supplied to the culture, incorporation into sn2-hydroxyarchaeol was observed after six days. Notably, the bacterial biomarkers were labelled more rapidly and substantially than the archaeal biomarker (Table 1). These results are comparable to those obtained for a similar ¹³C-labelling experiment with a consortium performing sulphate-dependent AOM²¹, in which rapid and substantial incorporation of ¹³C-labelled methane was noted in the lipids of sulphate-reducing bacteria, but minor ¹³C incorporation was observed in archaeal lipids only after prolonged incubation (>300 days). Our biomarker data thus indicate that a consortium consisting of an archaeon and a bacterium is responsible for AOM coupled to denitrification.

To determine the phylogenetic identity of the members of this consortium, we isolated genomic DNA from the biomass in the enrichment culture and constructed bacterial and archaeal 16S ribosomal RNA gene libraries. Sequence analysis of the bacterial clone library showed one dominant group of sequences clustering inside a subdivision with no other cultivated species (Fig. 2a and Supplementary Fig. 2a). This subdivision was distant from all other bacterial subdivisions (sequence identity less than 85%). Similar

Table 1 | Incorporation of methane into the bacterial and archaeal biomarkers of the enrichment culture

Biomarker compound	Amount (%)	Origin	Enrichment culture (t = 16 months)	After labelling with ¹³ C methane	
			δ ¹³ C (‰ versus VPDB)	3 days δ ¹³ C (‰ versus VPDB)	6 days δ ¹³ C (‰ versus VPDB)
Methane			-40.7		
Bicarbonate			-24.6		
C _{14:0} fatty acid	0.8	Bacteria	-26.1	+2,300	+4,400
Iso-C _{15:0} fatty acid	2.3	Bacteria	-31.0	+125	+189
C _{16:1} (Δ ⁹) fatty acid	6.8	Bacteria	-43.7	+4,300	-†
C _{16:0} fatty acid	11.4	Bacteria	-30.5	+370	+580
10-me-C _{16:1} (Δ ⁷) fatty acid	5.5	Bacteria	-37.9	-‡	-‡
10-me-C _{16:0} fatty acid	28.9	Bacteria	-38.5	-36.4	-36.5
C _{18:0} fatty acid	2.8	Bacteria	-31.4	+310	+380
C _{18:1} (Δ ¹¹ + Δ ¹⁰) fatty acid	14.3	Bacteria	-35.8	+900	-†
C ₁₉ cycloprop. fatty acid	12.7	Bacteria	-38.0	+290	+420
Diplopterol	4.1	Bacteria	-46.6	+380	-†
Sn2-hydroxyarchaeol	2.2	Archaea	-67.0	-76	-49

The table shows the relative abundances and stable carbon isotopic composition of the major lipids from the enrichment culture and the incorporation of ¹³C into these compounds after 3 or 6 days of incubation with ¹³C-labelled methane (δ¹³C = [(¹³C/¹²C)_{sample} / (¹³C/¹²C)_{standard}] - 1). For the archaeal compound from the enrichment culture, the large depletion in ¹³C indicates that its carbon is derived from ¹³C-depleted methane. The minor but significant incorporation of ¹³C into this compound during labelling (Δδ¹³C = -49 - (-67) = +18‰) suggests slow growth. VPDB, Vienna Pee Dee Belemnite. †δ¹³C after 6 days was not determined for some lipids already substantially enriched in ¹³C after 3 days. ‡δ¹³C could not be determined owing to low abundance of the compound and co-elution.

sequences from this subdivision have previously been retrieved from the denitrifying zone of sediments from Lake Biwa in Japan¹² and from contaminated groundwater in the United States¹³. The archaeal clone library contained a single sequence, which was only distantly related to the AOM Archaea of group 2 (ANME-2 (refs 9, 22), 86–87% identity) and cultivated methanogens (86–88% identity, Fig. 2b and Supplementary Fig. 2b). The highest similarity was found with archaeal clone sequences obtained from freshwater sediments from Lake Michigan in the United States¹⁴ and contaminated soils in Japan¹¹.

We used both the bacterial and archaeal 16S rRNA gene sequences to design specific probes for fluorescence *in situ* hybridization (FISH). Samples collected over the 16-month period of enrichment were now hybridized with these probes. In samples from the first three months, only occasional, single cells tested positive. Over time, both the bacterium and the archaeon became increasingly enriched, until they were the dominant microorganisms in the culture. After 16 months, approximately 10% of 4,6-diamidino-2-phenylindole (DAPI)-stained cells consisted of Archaea, all of which hybridized with the specific probe targeting the dominant archaeal sequence. The remainder of the culture consisted of bacteria, of which approximately 80% hybridized with the three specific probes targeting the dominant bacterial sequence (Fig. 2). As shown in Fig. 2, the Archaea were generally present as clusters inside a matrix of bacterial cells. The ratio of bacterial to archaeal cells (approximately 8:1) was different from the ratio reported for sulphate-dependent AOM¹⁸

(approximately 2:1). This difference might be explained by the higher energy yield of denitrification compared to sulphate reduction. α -, β - and γ -proteobacteria together made up <5% of the community. The sulphate reducers known to be involved in sulphate-dependent AOM¹⁰ were not detected, consistent with our observation that sulphate was not converted in the culture.

The nitrite-dependent AOM rate was 140 $\mu\text{mol CH}_4$ per g protein per hour (Fig. 1), corresponding to approximately 0.4 fmol CH_4 per cell per day for the Archaea in our enrichment culture. For sulphate-dependent AOM, a similar rate has been reported for the archaeal partner (0.7 fmol CH_4 per cell per day)¹⁸. This indicates that for AOM coupled to denitrification, the archaeal growth rate could be extremely low, with a doubling time in the order of several weeks, consistent with the long duration of the enrichment procedure.

To our knowledge this is the first report of archaeal AOM coupled to bacterial denitrification. In the 1970s, Mason⁸ discredited earlier studies that pure cultures of methanotrophic bacteria were able to denitrify using methane as the sole carbon and energy source. Instead, it was established experimentally that methanotrophs can oxidize methane aerobically to methanol or acetate at low oxygen concentrations, and that the methanol or acetate can subsequently be used to drive denitrification^{23–26}. Thus, the reaction presented here defines a new microbial guild with a potential contribution to biogeochemical cycling that has so far been overlooked. The recovery of related 16S rRNA gene sequences from different habitats and locations^{11–14} indicates that this process may contribute significantly

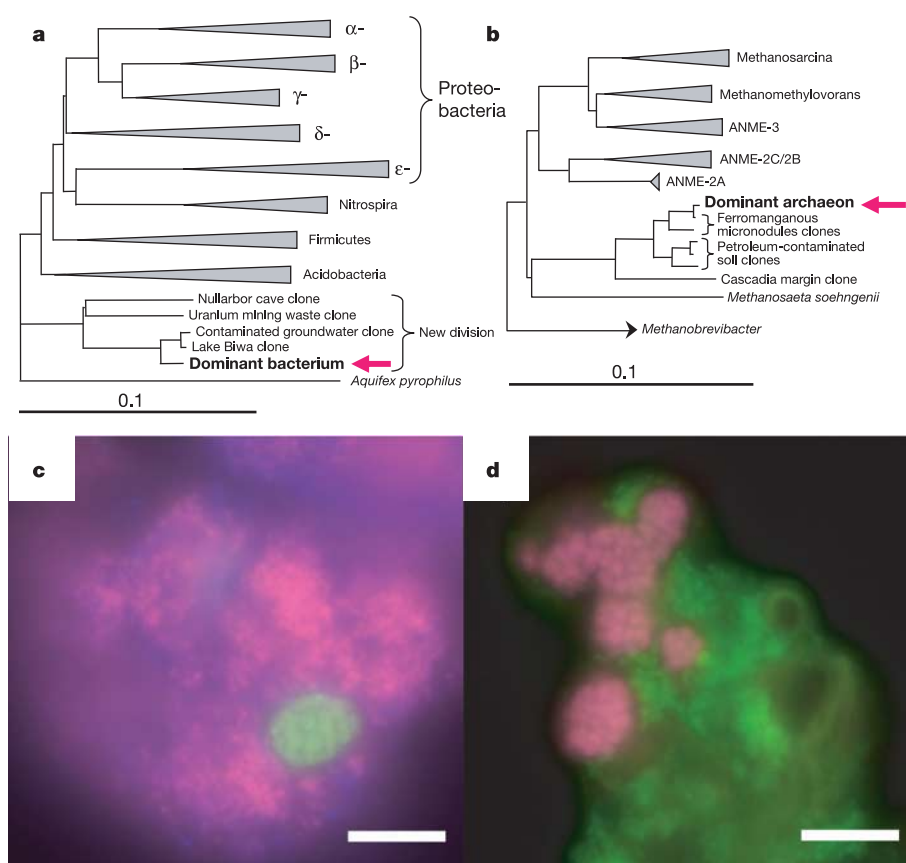


Figure 2 | Phylogeny and fluorescence *in situ* detection of the archaeal and bacterial members of the consortium mediating AOM coupled to denitrification. **a**, **b**, Consensus trees of the dominant bacterial (**a**) and archaeal (**b**) 16S rRNA sequences of the enrichment culture. Scale bars indicate 10 base substitutions per 100 bases. See also Supplementary Fig. 2 for more detailed trees, bootstrap values and treeing methods. **c**, Epifluorescence micrograph after hybridization with the general bacterial probe EUBmix (blue), the specific bacterial probe DBACT-193 (red) and the

specific archaeal probe DARCH-872 (green). The bacterial partner is pink, as it hybridizes with both the general and specific bacterial probes. **d**, Epifluorescence micrograph after hybridization with the general archaeal probe ARCH915 (blue), the specific archaeal probe DARCH-872 (red) and the general bacterial probe EUBmix (green). The archaeal partner is pink because it hybridizes with both the general and specific archaeal probes. Scale bars, 5 μm . See Methods for specification of probes.

to methane oxidation, and could potentially counteract worldwide increases in methane production associated with intensive agriculture. With biomarkers and probes for the responsible microorganisms now available, this possibility can be addressed.

METHODS

Sampling. Samples for inoculation were taken from the sediment of the Twentekanaal (52° 11' 04" N and 6° 24' 40" E, The Netherlands). Samples were collected from the top 15 cm of the sediment at 1 m water depth. At the time of sampling, the methane concentration at 15 cm sediment depth was 0.8 mM. The nitrate concentration in the water column was 0.1 mM. Cultivation and analytical methods are described in the Supplementary Information.

Calculations. The values of the Gibbs energy changes reported for equations (1) and (2) were calculated for standard conditions (25 °C, pH 7). The affinity constant of a microbial conversion is the substrate concentration at which the conversion rate is half of the maximum conversion rate. In this case the affinity constant for methane was estimated from the slope of the methane consumption in Fig. 1.

Methane incorporation and lipid analysis. Aliquots (60 ml) of the enrichment culture were anaerobically transferred to 120-ml serum bottles with an atmosphere of 90% argon and 10% ¹³C-labelled methane. The bottles were incubated on a shaker at 30 °C and then used for lipid analysis after three or six days. Lipids were ultrasonically extracted after freeze-drying and analysed by gas chromatography/mass spectrometry and isotope ratio gas chromatography/mass spectrometry as described previously²⁷. ¹³CO₂ was measured as the end product of AOM, using a GC-isotope ratio mass spectrometer (ThermoFinnigan Delta Plus).

16S rRNA gene sequence analysis and FISH. Chromosomal DNA from 1-ml reactor biomass was isolated and used as a template for polymerase chain reaction (PCR) amplification of 16S rRNA genes. PCR was performed using general bacterial primers (616F and 630R; ref. 27) and general archaeal primers (AR20F and AR958R; ref. 28). Cloning, sequencing and phylogenetic analyses were performed as described previously²⁷. On the basis of the bacterial and archaeal 16S rRNA gene sequences, new oligonucleotide probes were designed. The bacterial probes were S*-DBACT-0193-a-A-18 (5'-CGCTCGCCCC TTTGTC-3'), S*-DBACT-0447-a-A-18 (5'-CGCCGCCAAGTCATTCGT-3') and S*-DBACT-1027-a-A-18 (5'-TCTCCAGCTCCCTTGCG-3'), and the archaeal probe was S*-DARCH-0872-a-A-18 (5'-GGCTCCACCCGTTG TAGT-3'). We also used the general archaeal probe S-D-ARCH-0915-a-A-20, the general bacterial probe EUBmix and probe DSS658 for sulphate reducers¹⁰. FISH was performed as described previously²⁷. Formamide concentrations used in FISH experiments varied between 20% and 40%.

Received 30 November 2005; accepted 2 February 2006.

- Frankenberg, C., Meirink, J. F., Van Weele, M., Platt, U. & Wagner, T. Assessing methane emissions from global space-borne observations. *Science* **308**, 1010–1014 (2005).
- Galloway, J. N. *et al.* Nitrogen cycles: past, present, and future. *Biogeochemistry* **70**, 153–226 (2004).
- Conrad, R. Soil microorganisms as controllers of atmospheric trace gases (H₂, CO, CH₄, OCS, N₂O, and NO). *Microbiol. Rev.* **60**, 609–640 (1996).
- Knowles, R. Denitrifiers associated with methanotrophs and their potential impact on the nitrogen cycle. *Ecol. Eng.* **24**, 441–446 (2005).
- Shima, S. & Thauer, R. K. Methyl-coenzyme M reductase and the anaerobic oxidation of methane in methanotrophic Archaea. *Curr. Opin. Microbiol.* **8**, 643–648 (2005).
- Strous, M. & Jetten, M. S. M. Anaerobic oxidation of methane and ammonium. *Annu. Rev. Microbiol.* **58**, 99–117 (2004).
- Valentine, D. L. Biogeochemistry and microbial ecology of methane oxidation in anoxic environments: a review. *Anton. Leeuw. Int. J. G.* **81**, 271–282 (2002).
- Mason, I. Methane as a carbon source in biological denitrification. *J. Water Pollution Control Fed.* **49**, 855–857 (1977).
- Orphan, V. J. *et al.* Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. *Proc. Natl Acad. Sci. USA* **99**, 7663–7668 (2002).
- Boetius, A. *et al.* A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**, 623–626 (2000).
- Kasai, Y., Takahata, Y., Hoaki, T. & Watanabe, K. Physiological and molecular characterization of a microbial community established in unsaturated, petroleum-contaminated soil. *Environ. Microbiol.* **7**, 806–818 (2005).
- Koizumi, Y., Kojima, H. & Fukui, M. Characterization of depth-related microbial community structure in lake sediment by denaturing gradient gel electrophoresis of amplified 16S rDNA and reversely transcribed 16S rRNA fragments. *FEMS Microbiol. Ecol.* **46**, 147–157 (2003).
- Bakermans, C. & Madsen, E. L. Diversity of 16S rDNA and naphthalene dioxygenase genes from coal-tar-waste-contaminated aquifer waters. *Microb. Ecol.* **44**, 95–106 (2002).
- Stein, L. Y., La Duc, M. T., Grundl, T. J. & Neelson, K. H. Bacterial and archaeal populations associated with freshwater ferromanganese micronodules and sediments. *Environ. Microbiol.* **3**, 10–18 (2001).
- Hallam, S. J. *et al.* Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* **305**, 1457–1462 (2004).
- Krüger, M. *et al.* A conspicuous nickel protein in microbial mats that oxidize methane anaerobically. *Nature* **426**, 878–881 (2003).
- Strous, M., Kuenen, J. G., Fuerst, J. A., Wagner, M. & Jetten, M. S. The anammox case—a new experimental manifesto for microbiological eco-physiology. *Anton. Leeuw. Int. J. G.* **81**, 693–702 (2002).
- Nauhaus, K., Treude, T., Boetius, A. & Krüger, M. Environmental regulation of the anaerobic oxidation of methane: a comparison of ANME-I and ANME-II communities. *Appl. Environ. Microbiol.* **7**, 98–106 (2005).
- Koga, Y., Morii, H., Akagawa-Matsushita, M. & Ohga, M. Correlation of polar lipid composition with 16S rRNA phylogeny in methanogens. Further analysis of lipid component parts. *Biosci. Biotechnol. Biochem.* **62**, 230–236 (1998).
- Blumenberg, M., Seifert, R., Reitner, J., Pape, T. & Michaelis, W. Membrane lipid patterns typify distinct anaerobic methanotrophic consortia. *Proc. Natl Acad. Sci. USA* **101**, 11111–11116 (2004).
- Blumenberg, M., Seifert, R., Nauhaus, K., Pape, T. & Michaelis, W. *In vitro* study of lipid biosynthesis in an anaerobically methane-oxidizing microbial mat. *Appl. Environ. Microbiol.* **71**, 4345–4351 (2005).
- Knittel, K., Lösekann, T., Boetius, A., Kort, R. & Amann, R. I. Diversity and distribution of methanotrophic archaea at cold seeps. *Appl. Environ. Microbiol.* **71**, 467–479 (2005).
- Islas-Lima, S., Thalasso, F. & Gomez-Hernandez, J. Evidence of anoxic methane oxidation coupled to denitrification. *Water Res.* **38**, 13–16 (2004).
- Waki, M., Tanaka, Y., Osada, T. & Suzuki, K. Effects of nitrite and ammonium on methane-dependent denitrification. *Appl. Microbiol. Biotechnol.* **59**, 338–343 (2002).
- Eisentraeger, A., Klag, P., Vansbotter, B., Heymann, E. & Dott, W. Denitrification of groundwater with methane as sole hydrogen donor. *Water Res.* **35**, 2261–2267 (2001).
- Costa, C. *et al.* Denitrification with methane as electron donor in oxygen-limited bioreactors. *Appl. Microbiol. Biotechnol.* **53**, 754–762 (2000).
- Raghoebarsing, A. A. *et al.* Methanotrophic symbionts provide carbon for photosynthesis in peat bogs. *Nature* **436**, 1153–1156 (2005).
- Hinrichs, K. U., Hayes, J. M., Sylva, S. P., Brewer, P. G. & DeLong, E. F. Methane-consuming archaeobacteria in marine sediments. *Nature* **398**, 802–805 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank B. Kartal, J. van de Vossenberg and M. Schmid for discussions, and J. G. Kuenen and M. Wagner for critical reading of the manuscript. We thank J. Eigensteyn and W. Geerts for technical support. M.S. and K.F.E. are supported by a VIDI grant from the Dutch Science Foundation (NWO). Help from G. Boedeltje in choosing the sampling location is also gratefully acknowledged.

Author Contributions M.S., A.A.R., A.J.P.S. and K.T.P-S. performed the sampling; A.A.R. the enrichment, A.A.R., A.P. and K.F.E. the batch experiments and labelling; W.I.C.R., S.S. and J.S.S.D. the biomarker analysis; A.A.R. and K.T.P-S. the FISH analysis; A.A.R. the molecular analysis; and A.A.R. and H.J.M.O.C. the phylogeny and probe design. The research was conceived by M.S. and M.S.M.J., and pilot experiments were performed by K.T.P-S. A.A.R., A.P., K.F.E., S.S., J.S.S.D., H.J.M.O.C., M.S.M.J. and M.S. contributed to interpreting the data and writing the paper.

Author Information The 16S rRNA gene sequences have been deposited in GenBank under accession numbers DQ369741 (archaeal sequence) and DQ369742 (bacterial sequence). Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.S. (m.strous@science.ru.nl).

LETTERS

Nitrogen limitation constrains sustainability of ecosystem response to CO₂

Peter B. Reich¹, Sarah E. Hobbie², Tali Lee^{1,3}, David S. Ellsworth⁴, Jason B. West², David Tilman², Johannes M. H. Knops⁵, Shahid Naeem⁶ & Jared Trost¹

Enhanced plant biomass accumulation in response to elevated atmospheric CO₂ concentration could dampen the future rate of increase in CO₂ levels and associated climate warming. However, it is unknown whether CO₂-induced stimulation of plant growth and biomass accumulation will be sustained or whether limited nitrogen (N) availability constrains greater plant growth in a CO₂-enriched world^{1–9}. Here we show, after a six-year field study of perennial grassland species grown under ambient and elevated levels of CO₂ and N, that low availability of N progressively suppresses the positive response of plant biomass to elevated CO₂. Initially, the stimulation of total plant biomass by elevated CO₂ was no greater at enriched than at ambient N supply. After four to six years, however, elevated CO₂ stimulated plant biomass much less under ambient than enriched N supply. This response was consistent with the temporally divergent effects of elevated CO₂ on soil and plant N dynamics at differing levels of N supply. Our results indicate that variability in availability of soil N and deposition of atmospheric N are both likely to influence the response of plant biomass accumulation to elevated atmospheric CO₂. Given that limitations to productivity resulting from the insufficient availability of N are widespread in both unmanaged and managed vegetation^{5,7–9}, soil N supply is probably an important constraint on global terrestrial responses to elevated CO₂.

Continued emissions of CO₂ by the burning of fossil fuel are expected to increase global atmospheric CO₂ concentrations and surface temperatures significantly¹. These changes will occur much more rapidly if the terrestrial or oceanic carbon sinks weaken in the future. However, predicting future atmospheric CO₂ concentrations is hampered by uncertainties surrounding biospheric responses to elevated CO₂, such as the possibility that soil N availability restricts the magnitude of CO₂-induced biomass enhancement^{4–8}. However, the importance of, and the mechanisms underlying, this limitation to the CO₂ fertilization effect is a matter of debate^{2,3,7,8,10–12}. This is not surprising, because even with widespread N limitation of net primary production in terrestrial ecosystems¹³, the nature, timing and predictability of N constraints on ecosystem C accumulation remain unclear^{2,6,8,11}.

Multiple-resource-limitation theory provides a variety of mechanisms by which low N availability potentially limits the enhancement of biomass accumulation in response to elevated CO₂, but not all models agree that this will occur^{2,3,10,14,15}. At elevated CO₂, low supply of N in the soil could limit the leaf area for intercepting light, limit the capacity of plants to fix CO₂ photosynthetically, or lead to diminishing plant N availability over time through long-term soil-plant C and N dynamics^{2–4,8,11,15,16}. Each of these, or their combination, could diminish biomass response to elevated CO₂. Therefore,

a large increase in biomass accumulation under elevated CO₂, often observed in short-term experiments, may not be sustained over the long term in natural systems given the N limiting conditions that predominate in both unmanaged and managed vegetation^{5,7,8}.

Much research has addressed the response of N availability to elevated CO₂, but findings have been inconsistent. Soil N availability has been observed to decrease under elevated CO₂, but it has also been observed to remain stable or increase^{8,12,16–19}. Moreover, results from studies that manipulated both CO₂ and N are variable in terms of whether biomass accrual under elevated CO₂ is affected by N supply^{5,6,9,18,20–22}. The inconsistency of such responses may reflect inconsistent responses of soil N availability to elevated CO₂, but additionally they may reflect variation in nutrient-use efficiency, stoichiometry or soil microbial processes that may eliminate or delay the impact of N limitation on the elevated CO₂ response^{2,8,12}.

Most experiments that have manipulated CO₂ and N^{5,6,18,20–22} have been short in duration (3 years or less), which limits their utility in understanding interactions that are predicted to occur over longer time frames^{2,3,8,15}. Despite the importance of understanding CO₂ × N interactions, to our knowledge there are only three replicated field experiments in the world that have manipulated both CO₂ and N supply for more than three years (including ours reported here). Two of these experiments are relevant to natural systems (early results are reported in refs 6, 23) and one to intensively managed agricultural systems (long-term results are reported in ref. 9). In our experiment (called BioCON) we grew 296 field plots containing different numbers (1, 4, 9 or 16 species) and combinations of perennial grassland species under ambient and elevated (560 μmol mol^{−1}) atmospheric CO₂ and with either ambient or enriched soil N supply (ambient and enriched with 4 g N m^{−2} yr^{−1} as NH₄NO₃). A total of 16 native or naturalized species were used in the study, four each from four functional groups (legumes, non-leguminous forbs, C₃ grasses and C₄ grasses). The main effects of diversity and interactions with CO₂ and N were addressed previously^{18,23}. We proposed that biomass accumulation of plants growing in ambient, N-poor soil would be less responsive to elevated CO₂ than would biomass of plants grown in N-enriched soil. The contrasting high versus low levels of N supply in our study are broadly consistent with the global range in both rates of N supply in unfertilized soils and in annual N deposition^{13,24}.

During the six years of the study, the stimulation of plant biomass accumulation by elevated CO₂ became continuously weaker under ambient N supply than under enriched N supply (Fig. 1). Both CO₂ and N generally stimulated biomass accumulation (Fig. 1, Table 1), and overall biomass levels declined after the first year and then varied over time. However, for the entire period there was a significant

¹Department of Forest Resources, ²Department of Ecology, Evolution and Behavior, University of Minnesota, St Paul, Minnesota 55108, USA. ³Department of Biology, University of Wisconsin-Eau Claire, Eau Claire, Wisconsin 54701, USA. ⁴School of Natural Resources and Environment, University of Michigan, Ann Arbor, Michigan 48109, USA. ⁵School of Biological Sciences, University of Nebraska, Lincoln, Nebraska 68588, USA. ⁶Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, New York 10027, USA.

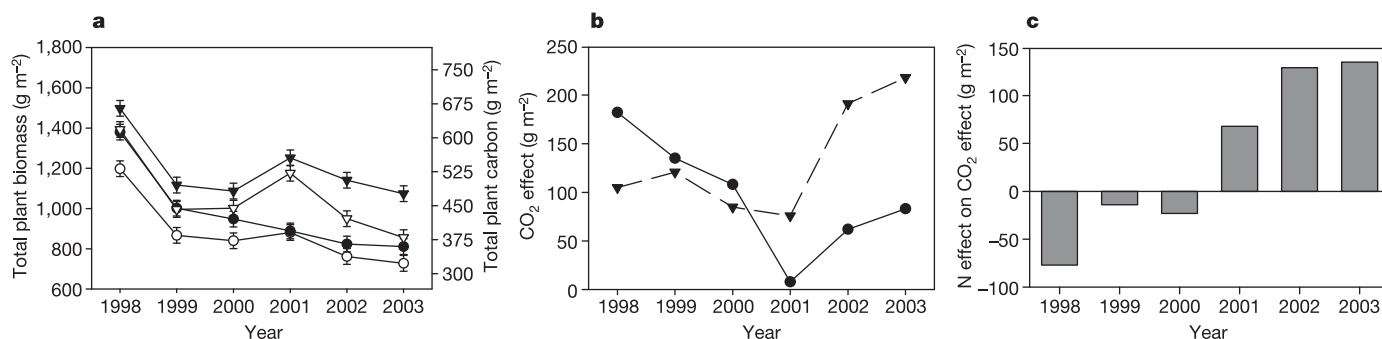


Figure 1 | Effects of CO₂ and N on total plant biomass over time. **a**, Total plant biomass (above-ground plus 0–20 cm below-ground) and carbon at ambient and elevated CO₂ × N levels from 1998 to 2003. Data were pooled across diversity treatments, and each point shows the annual mean (plus standard error) of 74 plots sampled twice per year. There was a significant interaction (Table 1) between CO₂, N and year ($P = 0.0013$), plus significant main effects ($P < 0.05$) of year, CO₂ and N. Open circles, ambient CO₂ and ambient N; filled circles, elevated CO₂ and ambient N; open triangles,

ambient CO₂ and enriched N; filled triangles, elevated CO₂ and enriched N. **b**, The effect of CO₂ on total biomass: the CO₂ enhancement (assessed as (value at elevated CO₂) minus (value at ambient CO₂)) at ambient (circles) and enriched (triangles) N supply each year. **c**, The effect of N availability on the CO₂ biomass effect, defined as the difference between the CO₂ effect at enriched N and that at ambient N, namely (CO₂ effect at enriched N) minus (CO₂ effect at ambient N).

temporal shift in the enhancement of total biomass by CO₂ at ambient versus enriched N ($P = 0.0013$ for the CO₂ × N × year interaction; Table 1, Fig. 1). This impact over time of N supply rate on response to elevated CO₂ did not vary significantly with species diversity treatment. Data are therefore shown pooled across all species richness levels.

The significant three-way interaction involving CO₂, N and year occurred because the influence of enriched N supply on the CO₂ response shifted in direction and grew larger with time (Fig. 1b, c, Table 1). The stimulation of total biomass by elevated CO₂ was higher at ambient N than at enriched N each year from 1998 to 2000 (Fig. 1b), although for each year analysed individually there was no significant effect of N on the elevated CO₂ response. However, in 2001, 2002 and 2003 the sign of the CO₂ × N interaction switched: stimulation of total biomass by CO₂ was greater at enriched N supply than at ambient N supply (statistically significant CO₂ × N interaction in 2002 and 2003; $P < 0.05$). For example, in 2002 and 2003, respectively, elevated CO₂ increased biomass by 62 and 82 g m⁻² at ambient N and by 191 and 219 g m⁻² at enriched N (Fig. 1b). These correspond to 20–25% stimulation of biomass by elevated CO₂ at enriched N, compared with 8–12% stimulation at ambient N.

The mechanisms responsible for the changes with time in the response of biomass accumulation to CO₂ and N supply (Fig. 1) probably involve temporal heterogeneity in physiological acclimation, soil C and N mineralization and immobilization, and litter, root and soil organic matter decomposition and turnover^{12,18,25–28}. For example, there were temporal changes in CO₂ × N effects on total plant N pools and on net N mineralization rates in soil, including significant ($P < 0.05$) interactions between CO₂, N and year in both cases (Table 1). These were a consequence of elevated CO₂ substantially increasing plant N pools and soil net N mineralization rates at enriched N supply in 2002 and 2003 (Supplementary Figs 1 and 2). These biomass and N cycle shifts in CO₂ × N response over time could result in part from the immobilization of N into living tissues (below-ground turnover is about three years) and decomposing organic matter. Such N sequestration would probably have been enhanced in the treatment with elevated relative to ambient CO₂ in the first two years of the study, because of higher biomass, initially higher biomass C:N ratios, and greater labile C inputs²⁸. The positive response to CO₂ enrichment observed under ambient N during the first two years of the experiment may also be more characteristic of seedlings or juvenile plants than of mature, well-established plants.

Additionally, to test whether the accumulation of greater plant biomass at elevated CO₂ than at ambient CO₂ was due to a greater efficiency of use of N pools in C acquisition and accumulation, or to

greater N pools, we examined the relationship between total plant biomass and total plant N across all plots. The slope of this relationship was greater at elevated CO₂ than at ambient CO₂ in 1998 and 1999 but not in 2000–2003 (Supplementary Table 1 and Supplementary Fig. 3; CO₂ treatment × plant N pool × year interaction, $P < 0.0001$). In 1998–1999, total plant biomass was about 15% higher at any given plant N pool at elevated CO₂ than at ambient CO₂. This CO₂-enhanced N use efficiency disappeared with time: from 2000 to 2003 plots with elevated CO₂ tended to occupy positions higher up on the same biomass–N relationship. These results indicate that in 1998–1999, stimulation of biomass accumulation by elevated CO₂ might have been related to increased efficiency of physiological N use, whereas in later years the CO₂ stimulation of biomass was associated with shifts in total plant N availability and accumulation.

We conclude from the results of our study that long-term biomass accumulation was influenced by the relative availability of CO₂ and N, with ambient N supply gradually limiting the potential biomass accumulation in response to elevated CO₂, and enriched N supply gradually stimulating this response. These responses were similar for simple monocultures and complex mixtures. Moreover, even the plots with legumes, generally considered responsive to elevated CO₂ independently of N supply^{26,29,30}, did not respond differently in terms of CO₂ × N interactions than plots without legumes. In separate analyses of variance there was no significant interaction of CO₂ × N × year with the presence or abundance of legumes or *Lupinus* (the most abundant legume) (Supplementary Table 2). Thus, the presence of legumes did not prevent the gradual development of a diminished elevated CO₂ biomass response at ambient N³⁰.

Our results are consistent with and extend previous experimental studies of CO₂ × N interactions in agricultural and forest plantation systems^{5,9}. With its wide range of species types and species combinations, including mixtures, our study provides a broad test of CO₂ × N interactions under contrasting but relatively low N supply rates, and includes measurements of root biomass, whereas previous studies have been in a limited range of species monocultures with higher rates of N addition, and included no below-ground biomass measures^{5,9}. In a two-year manipulation of CO₂ × N in a 15-year-old pine plantation, wood production increased with elevated CO₂ with a N addition of 11 g N m⁻² yr⁻¹ but not under ambient N conditions⁵. In a long-term study in a highly managed pasture⁹, harvestable above-ground biomass in *Lolium* monocultures increased by more than 150 g m⁻² yr⁻¹ with elevated CO₂ when N addition was 56 g N m⁻² yr⁻¹ but did not increase when N addition was 14 g N m⁻² yr⁻¹.

Table 1 | Repeated-measures analysis of variance of CO₂ and N effects

Effect	Total biomass (g m ⁻²)		Total plant N (g m ⁻²)		Net soil N mineralization (mg kg ⁻¹ d ⁻¹)	
	F	P > F	F	P > F	F	P > F
CO ₂	11.17	0.0288	52.95	0.1609	0.01	0.9303
N	46.77	<0.0001	77.90	<0.0001	1.55	0.2141
Year	471.52	<0.0001	1.87	0.1714	7.06	0.0080
CO ₂ × year	0.01	0.9283	11.68	0.0006	3.64	0.0566
N × year	9.84	0.0017	1.46	0.2264	1.23	0.2671
CO ₂ × N	0.38	0.5356	0.01	0.9196	1.10	0.2963
CO ₂ × N × year	10.35	0.0013	3.86	0.0497	4.25	0.0394
Diversity	133.25	<0.0001	75.57	<0.0001	16.21	<0.0001
Whole-model R ²	0.58	<0.0001	0.40	<0.0001	0.21	<0.0001

Effects of year, CO₂, N and diversity, and all interactions, on total biomass (above-ground plus 0–20 cm below-ground), total plant N per unit ground area, and net soil N mineralization rate are shown. Year is included in the model as a continuous variable. Diversity was a significant main effect in all analyses but it did not influence the CO₂ × N interaction (CO₂ × N × diversity, $P \geq 0.4$) or the year × CO₂ × N interaction (CO₂ × N × year × diversity, $P \geq 0.4$). Interaction terms including diversity effects are therefore not shown above, and effects of year, CO₂ and N are shown pooled across diversity treatments in Fig. 1 and Supplementary Information. Significant effects ($P \leq 0.05$) are shown in bold text.

Findings from both of these studies are consistent with our results, indicating that the overarching control on responses to elevated CO₂ by C:N supply stoichiometry might be comparable in intensive agricultural, forest plantation and natural ecosystems, although the N levels required to alleviate N limitation in agricultural systems may be an order of magnitude higher than in unmanaged vegetation. Given that the levels in our N treatments (ambient and +4 g N m⁻² yr⁻¹), and their difference, were within the global range of variation of both natural soil N availability and of atmospheric N deposition^{13,24}, the results indicate that variabilities in soil N and in atmospheric N deposition are both likely to influence terrestrial ecosystem responses to elevated CO₂. In sites with low-to-moderate soil N availability, N limitation probably constrains the stimulation of biomass accumulation by elevated CO₂. In soils that are N-rich and in regions with very high levels of N deposition, such limitation of the CO₂ fertilization effect by insufficient N may be weak or absent.

Our results are consistent with the idea⁷ that although some models indicate a considerable capacity of land ecosystems to sequester large amounts of C in the coming century¹, this C accumulation is likely to be constrained over time by N availability^{7,8}. Moreover, predictions of global C and N dynamics, and policy and political decisions whose efficacy rests on such predictions, should incorporate the impacts of multiple global change factors such as anthropogenic fertilization with C and N, as well as the interactions between these and other global factors. Estimating the role of terrestrial ecosystems as current and future carbon sinks hinges on the ability to understand the biogeochemical consequences of variability in soil N limitations and N deposition rates worldwide.

METHODS

In our experiment (called BioCON) we grew 296 field plots (each 2 m × 2 m) containing different numbers and combinations of perennial grassland species under ambient and elevated atmospheric CO₂ and with either ambient or enriched soil N supply^{18,25}. Plots were arranged in six circular 20-m diameter rings at the Cedar Creek Natural History Area in central Minnesota, USA. In three elevated-CO₂ rings, a free-air CO₂ enrichment system¹⁸ was used during each growing season to maintain the CO₂ concentration at an average of 560 μmol mol⁻¹, a concentration likely to be reached this century. Three ambient CO₂ rings were treated identically but without additional CO₂. Half of the plots in each ring received N amendments of 4 g N m⁻² yr⁻¹ applied as NH₄NO₃ on three dates each year. The treatments were arranged in complete factorial combination of two levels of atmospheric CO₂ (ambient and elevated), four levels of plant species richness (1, 4, 9 and 16) and two levels of N (ambient and enriched with added N) (see Supplementary Information and refs 18, 25 for

further detail). There were 32, 15, 15 and 12 plots with 1, 4, 9 and 16 species, respectively, at each of the four contrasting CO₂ and N levels ($n = 296$ plots in all). The 16 study species included four C₄ grasses (*Andropogon gerardii*, *Bouteloua gracilis*, *Schizachyrium scoparium* and *Sorghastrum nutans*), four C₃ grasses (*Agropyron repens*, *Bromus inermis*, *Koeleria cristata* and *Poa pratensis*), four N-fixing legumes (*Amorpha canescens*, *Lespedeza capitata*, *Lupinus perennis* and *Petalostemum villosus*) and four non-leguminous forb species (*Achillea millefolium*, *Anemone cylindrica*, *Asclepias tuberosa* and *Solidago rigida*).

Each year in every plot we assessed both above-ground and below-ground (0–20 cm) plant biomass (in both June and August) and plant C and N concentrations (in August only)^{18,25}. Above-ground biomass was harvested by clipping a 10 cm × 100 cm strip just above the soil surface. Total below-ground biomass (fine roots, coarse roots and crowns) was sampled at 0–20 cm depth with the use of three cores 5 cm in diameter. Soil net N mineralization rates were measured *in situ* each year in each plot by using a semi-open core one-month incubation¹⁸ beginning in late June. We examined results for the entire period 1998–2003 and used a repeated-measures analysis of variance to test for the main effects of CO₂, N and diversity and their interactions, and whether these changed with year.

Received 10 October; accepted 29 November 2005.

- Intergovernmental Panel on Climate Change. *Climate Change 2001: The Scientific Basis* (Cambridge Univ. Press, New York, 2001).
- Rastetter, E. B., Agren, G. I. & Shaver, G. R. Responses of N-limited ecosystems to increased CO₂: a balanced-nutrition coupled-element-cycles model. *Ecol. Appl.* **7**, 444–460 (1997).
- McMurtrie, R. E., Dewar, R. C., Medlyn, B. E. & Jeffreys, M. P. Effects of rising CO₂ on growth and carbon sequestration in forests: a modelling analysis of the consequences of altered litter quantity and quality. *Plant Soil* **224**, 135–152 (2000).
- Zak, D. R., Pregitzer, K. S., Curtis, P. S. & Holmes, W. E. Atmospheric CO₂ and the composition and function of soil microbial communities. *Ecol. Appl.* **10**, 47–59 (2000).
- Oren, R. *et al.* Soil fertility limits carbon sequestration by forest ecosystems in a CO₂ enriched world. *Nature* **411**, 469–472 (2001).
- Shaw, M. R. *et al.* Grassland responses to global environmental changes suppressed by elevated CO₂. *Science* **298**, 1987–1990 (2002).
- Hungate, B. A., Dukes, J. S., Shaw, M. R., Luo, Y. Q. & Field, C. B. Nitrogen and climate change. *Science* **302**, 1512–1513 (2003).
- Luo, Y. *et al.* Progressive nitrogen limitation of ecosystem responses to rising atmospheric carbon dioxide. *Bioscience* **54**, 731–739 (2004).
- Schneider, M. K. *et al.* Ten years of free-air CO₂ enrichment altered the mobilization of N from soil in *Lolium perenne* L. swards. *Glob. Change Biol.* **10**, 1377–1388 (2004).
- Lloyd, J. & Farquhar, G. D. The CO₂ dependence of photosynthesis, plant growth responses to elevated atmospheric CO₂ concentrations, and their interactions with soil nutrient status. I. General principles and forest ecosystems. *Funct. Ecol.* **10**, 4–32 (1996).
- Peterson, A. G. *et al.* The photosynthesis–leaf nitrogen relationship at ambient and elevated atmospheric carbon dioxide: a meta-analysis. *Glob. Change Biol.* **5**, 331–346 (1999).
- Zak, D. R., Holmes, W. E., Finzi, A. C., Norby, R. J. & Schlesinger, W. H. Soil nitrogen cycling under elevated CO₂: a synthesis of forest face experiments. *Ecol. Appl.* **13**, 1508–1514 (2003).
- Vitousek, P. M. *et al.* Human alteration of the global nitrogen cycle: sources and consequences. *Ecol. Appl.* **7**, 737–750 (1997).
- Luo, Y., Field, C. B. & Mooney, H. A. Predicting responses of photosynthesis and root fraction to elevated [CO₂]: interactions among carbon, nitrogen and growth. *Plant Cell Environ.* **17**, 1195–1204 (1994).
- Cannell, M. & Thornley, J. N-poor ecosystems may respond more to elevated [CO₂] than N-rich ones in the long term. A model analysis of grassland. *Glob. Change Biol.* **4**, 431–442 (1998).
- Gill, R. A. *et al.* Nonlinear grassland responses to past and future atmospheric CO₂. *Nature* **417**, 279–282 (2002).
- Williams, M. A., Rice, C. W. & Owensby, C. E. Carbon dynamics and microbial activity in tallgrass prairie exposed to elevated CO₂ for 8 years. *Plant Soil* **227**, 127–137 (2000).
- Reich, P. B. *et al.* Plant diversity enhances ecosystem responses to elevated CO₂ and nitrogen deposition. *Nature* **410**, 809–812 (2001).
- Finzi, A. C. & Schlesinger, W. H. Soil-nitrogen cycling in a pine forest exposed to 5 years of elevated carbon dioxide. *Ecosystems* **6**, 444–456 (2003).
- Owensby, C. E., Auen, L. M. & Coyne, P. I. Biomass production in a nitrogen fertilized, tallgrass prairie ecosystem exposed to ambient and elevated levels of CO₂. *Plant Soil* **165**, 105–113 (1994).
- Zanetti, S. *et al.* Does nitrogen nutrition restrict the CO₂ response of fertile grassland lacking legumes? *Oecologia* **112**, 17–25 (1997).
- Zak, D. R., Pregitzer, K. S., Curtis, P. S. & Holmes, W. E. Atmospheric CO₂ and the composition and function of soil microbial communities. *Ecol. Appl.* **10**, 47–59 (2000).

23. Reich, P. B. *et al.* Species and functional diversity independently influence biomass accumulation and its response to CO₂ and N. *Proc. Natl Acad. Sci. USA* **101**, 10101–10106 (2004).
24. Galloway, J. N. *et al.* The nitrogen cascade. *Bioscience* **53**, 341–356 (2003).
25. Reich, P. B. *et al.* Do species and functional groups differ in acquisition and use of C, N and water under varying atmospheric CO₂ and N availability regimes? A field test with 16 grassland species. *New Phytol.* **150**, 435–448 (2001).
26. Lee, T. D., Tjoelker, M. G., Ellsworth, D. S. & Reich, P. B. Leaf gas exchange responses of 13 prairie grassland species in the field under elevated carbon dioxide and increased nitrogen supply. *New Phytol.* **150**, 405–418 (2001).
27. Craine, J. M., Wedin, D. A. & Reich, P. B. The response of soil CO₂ flux to changes in atmospheric CO₂, nitrogen supply, and plant diversity. *Glob. Change Biol.* **7**, 947–953 (2001).
28. Dijkstra, F., Hobbie, S., Reich, P. & Knops, J. Divergent effects of elevated CO₂, N fertilization, and plant diversity on soil C and N dynamics in a grassland field experiment. *Plant Soil* **272**, 41–52 (2005).
29. Lüscher, A., Hartwig, U. A., Suter, D. & Nösberger, J. Direct evidence that in fertile grassland symbiotic N₂ fixation is an important trait for a strong response of plants to elevated atmospheric CO₂. *Glob. Change Biol.* **6**, 655–662 (2000).
30. Hungate, B. A. *et al.* CO₂ elicits long-term decline in nitrogen fixation. *Science* **304**, 1291 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the US Department of Energy Program for Ecosystem Research, the National Science Foundation Long-Term Ecological Research and Biocomplexity Coupled Biogeochemical Cycles Programs, and the University of Minnesota for supporting this research.

Author Contributions All authors were involved in the design or implementation of the experiment, or both, with P.B.R. having the major responsibility for these. P.B.R., D.E., D.T., J.K. and S.N. were involved in the initial experimental design. J.T. was responsible for a substantial fraction of the experimental management and data collection, with smaller contributions in this area from T.L., D.E., J.B.W., J.K. and S.H. P.B.R. performed the statistical analyses and wrote the paper, with assistance from S.H., T.L., D.E., S.N. and J.B.W. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.B.R. (preich@umn.edu).

LETTERS

Parental investment by skin feeding in a caecilian amphibian

Alexander Kupfer¹, Hendrik Müller^{1,2}, Marta M. Antoniazzi³, Carlos Jared³, Hartmut Greven⁴, Ronald A. Nussbaum⁵ & Mark Wilkinson¹

Although the initial growth and development of most multicellular animals depends on the provision of yolk, there are many varied contrivances by which animals provide additional or alternative investment in their offspring¹. Providing offspring with additional nutrition should be favoured by natural selection when the consequent increased fitness of the young offsets any corresponding reduction in fecundity². Alternative forms of nutrition may allow parents to delay and potentially redirect their investment. Here we report a remarkable form of parental care and mechanism of parent–offspring nutrient transfer in a caecilian amphibian. *Boulengerula taitanus* is a direct-developing, oviparous caecilian³, the skin of which is transformed in brooding females to provide a rich supply of nutrients for the developing offspring. Young animals are equipped with a specialized dentition, which they use to peel and eat the outer layer of their mother's modified skin. This new form of parental care provides a plausible intermediate stage in the evolution of viviparity in caecilians. At independence, offspring of viviparous and of oviparous dermatotrophic caecilians are relatively large despite being provided with relatively little yolk. The specialized dentition of skin-feeding (dermatophagous) caecilians may constitute a pre-adaptation to the fetal feeding on the oviduct lining of viviparous caecilians.

Amphibians are renowned for their diverse forms of parental investment, including hiding, guarding, transporting or feeding their offspring^{4,5}. The reproductive diversity of the tropical caecilian amphibians is more poorly known than that of salamanders and frogs, although it is known to include both oviparity, with an aquatic larva or direct development, and viviparity. Viviparous caecilians are unusual in having a specialized, deciduous, fetal dentition^{6,7} which is thought to be used to scrape lipid-rich secretions and cellular materials from the hypertrophied lining of the maternal oviduct^{6,8–11}. In contrast, it is generally thought that oviparous caecilians provision their offspring only with yolk, with additional investment limited to the attendance of egg clutches¹².

In the 1990s we discovered teeth in hatchlings of the oviparous Neotropical caecilian *Siphonops annulatus* that resemble more the fetal teeth of viviparous caecilians than the teeth of adults of this species¹³. Field observations revealed that hatchlings are altricial and remain with their mothers until they have grown substantially. Mothers also have a paler skin colour than non-attending adults. Speculating on these observations, we proposed that the fetal-like dentition of newborn *S. annulatus* is used to feed on glandular secretions of the mother's skin¹⁴, in a manner analogous to that of mammalian lactation. Here we report detailed observations of *Boulengerula taitanus*, another oviparous³ caecilian species that has

altricial¹⁵ young (see Fig. 1a) equipped with 'fetal-like' teeth, including observations of several bouts of feeding. Rather than scraping up skin secretions, the young of *B. taitanus* use their teeth to peel and eat the specially modified skin of their mothers.

Twenty-one females of the Kenyan caecilian *B. taitanus*, with broods of between two and nine young, were collected from subterranean nest sites and maintained and observed in captivity. Eight episodes of skin feeding by different young from five different broods were observed, and five were filmed (see Fig. 1b, and Supplementary movies 1 and 2). In each episode, the young moved over and around their mother's bodies, vigorously pressing their heads against their mothers while repeatedly opening and closing their mouths, and using their lower jaws in particular to lift and peel the outer layer of the mother's skin.

During one week of maternal care the young increased their total length substantially (about 11%; Fig. 1c) with average individual growth estimated to be about 1 mm per day. No alternative feeding of

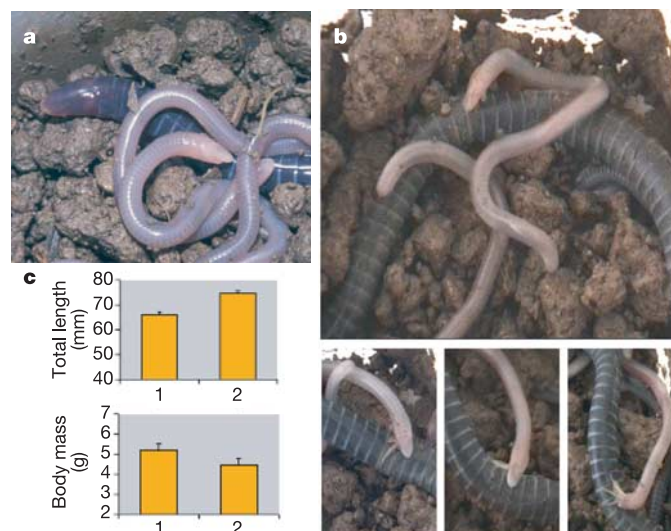


Figure 1 | Skin feeding in *B. taitanus*. **a**, Female with unpigmented young. **b**, Various stills from video footage of a young animal peeling and eating the outermost layer of its mother's skin. **c**, Changes in mean total length ($n = 66$, $P < 0.001$; t -test) of young (top) and mean body mass ($n = 15$, $P < 0.001$; paired t -test) of mothers (bottom) between a first (1) and a second (2) measurement after one week of parental care. Error bars show s.e.m.

¹Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK. ²Institute of Biology, Leiden University, Kaiserstraat 63, 2311 GP, Leiden, The Netherlands. ³Laboratório de Biologia Celular, Instituto Butantan, Av. Vital Brasil 1500, 05503-900, São Paulo, Brazil. ⁴Institut für Zoomorphologie und Zellbiologie der Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse, D-40225 Düsseldorf, Germany. ⁵Museum of Zoology and Department of Ecology and Evolutionary Biology, University of Michigan Ann Arbor, Michigan 48109, USA.

young was observed, and the stomachs of control young preserved in the field immediately on collection contained only monolayers of skin, indicating that ingested skin alone provides sufficient nutrients for the considerable growth observed. Maternal weight loss over the same period (about 14%; Fig. 1c) is consistent with skin feeding imposing a high cost on mothers.

Most attending females of *B. taitanus* were notably paler than non-attending adults, reflecting differences at cellular and tissue levels associated with the skin's role in nutrition (Fig. 2). The outermost layer of the epidermis, the stratum corneum, typically comprises squamous (flattened), keratinized cells (Fig. 2a), whereas in brooding females the cells are far more voluminous and full of vesicles (Fig. 2b). Overall, the epidermis of brooding females is up to twice the thickness of that of non-brooding females, as a result of elongation of the stratified epithelial cells rather than any increase in numbers of cells. Histochemistry reveals that, unlike in non-brooding females, the cytoplasm of modified epidermal cells of brooding females is full of lipid inclusions (staining positive for sudan black B). Tests for carbohydrate (alcian blue and periodic acid–Schiff) were negative. The stratum corneum is also rich in protein (staining positive for bromophenol blue) in both brooding and non-brooding females.

Adult *B. taitanus* are predators and have two rows of pointed teeth in the upper (premaxillary–maxillary and vomeropalatine) and in the lower (dentary and splenial) jaws, with either one (Fig. 3a) or two (a labial and more apical lingual) distinct cusps (Fig. 3b)³. Whereas tooth crowns of larval indirect-developing caecilians resemble those

of adults, those of the dermatophagous young of *B. taitanus* are remarkably divergent and variable (Fig. 3c–f). Bicuspid splenial teeth are present but not yet erupted. The vomeropalatine teeth and the anteriormost three or four teeth of the premaxillary–maxillary and the dentary series are monocuspid. The remaining teeth are multicuspid, combining a pronounced blade-like labial cusp with a lingual cusp that has two or three subsidiary cusps (Fig. 3d), which may be short and blunt (Fig. 3e) or have more elongate and pointed processes resembling a grappling hook (Fig. 3f). Hatchling *B. taitanus* (total length about 28 mm) have several other unusual characteristics, seemingly associated with their altriciality. The skull and axial skeleton are mostly poorly ossified in comparison with hatchlings of direct-developing species¹⁶, and the body musculature and associated external annulation are weakly developed, severely constraining mobility. By the time they become independent of their mothers

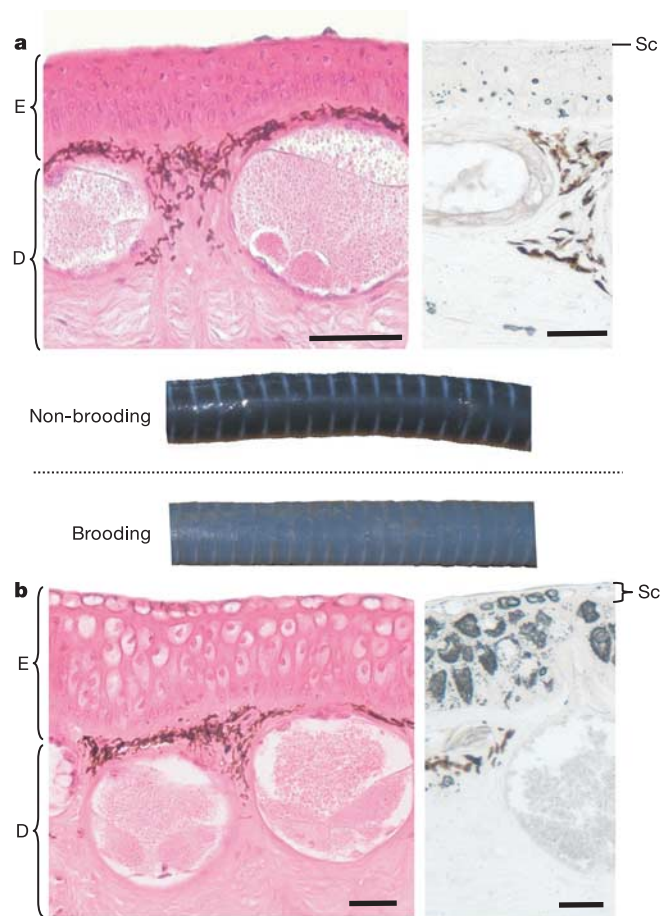


Figure 2 | Comparison of the skin of non-brooding and brooding female *B. taitanus*, showing differences in skin colour, structure and histochemistry. a, Non-brooding; b, brooding. Sections on the left were stained with haematoxylin and eosin; those on the right were stained (dark) for lipids with sudan black B. E, epidermis; D, dermis; Sc, stratum corneum. Scale bars, 50 µm.

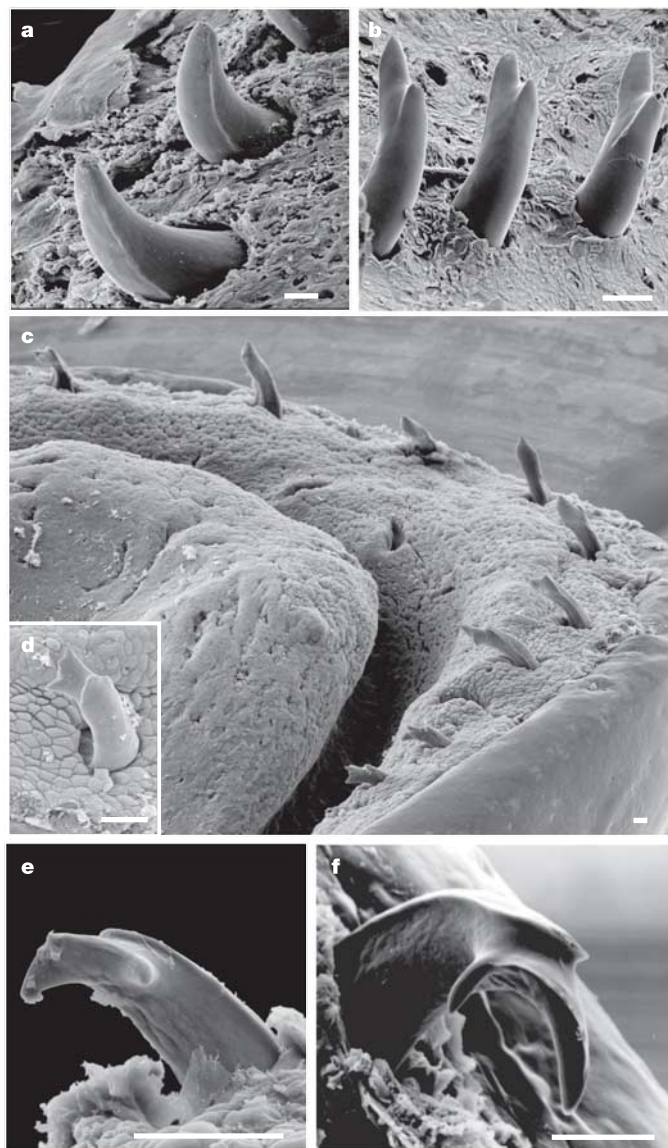


Figure 3 | Dentition of adult and young *B. taitanus*. a, Anterior view of two monocuspid, adult premaxillary teeth. b, Labial view of three bicuspid, adult vomerine teeth. c, Lateral view of a lower jaw of a young specimen (total length 69 mm), showing different dentary tooth crown morphologies. d, Labial view of a posterior dentary tooth of this young specimen. e, Anterior premaxillary tooth of the same specimen. f, Anterior premaxillary tooth of a smaller specimen (total length 57 mm) resembling a grappling hook. Scale bars, 30 µm.

(total length about 86 mm), young resemble miniature adults in these features.

Dermatophagy, as seen in *B. taitanus*, is a highly unusual mode of parental care previously unknown in tetrapods, in which nutrient provisioning involves remarkable adaptations of both the mothers and the young. Many vertebrates periodically shed their stratum corneum and some, including caecilians, eat and recycle nutrients from their own shed skin (autodermatophagy)¹⁷. In contrast, the altricial young of *B. taitanus* depend for a time entirely on their mother's skin, which is suitably transformed to provide nutrient that, like mammalian milk, is rich in lipids. Amphibian skin is well known for its diverse functions^{18,19}, and its role in *B. taitanus* can be presumed to impose constraints upon other normal functions. For example, dermal granular glands are frequently associated with toxic secretions with a defensive function in amphibians¹⁹, and some downregulation of toxins during skin feeding might be expected. Aggressive dermatophagy could injure the mother, and we might also expect the periodic bouts of feeding to be more or less synchronized with the maternal sloughing cycle, which may itself be modified, and to involve some signalling between parent and offspring. There is clearly scope for both parent-offspring conflict and sibling competition where there is dermatotrophic parental care.

Oviparous caecilians were previously believed only to guard their eggs until hatching and to provide no subsequent parental care^{8,12}, similarly to *Ichthyophis*²⁰. This nutritional investment in offspring only in the form of yolk (lecithotrophy) is seen in all 'primitive' caecilians and is inferred to be the ancestral condition, with viviparity, and fetal feeding on the maternal oviduct lining (matrotrophy), being derived. Maternal dermatotrophy provides a highly plausible intermediate between these different reproductive modes. Current understanding of caecilian phylogeny²¹ indicates that viviparity must have evolved independently several times in caecilians, which implies striking and enigmatic convergent evolution of the associated fetal teeth. The discovery of fetal-like teeth in maternal dermatotrophic caecilians indicates that although viviparity is convergent in caecilians, one of its most distinctive features, fetal teeth, might not be. Fetal-like teeth are known also in some species of the oviparous Neotropical genera *Siphonops* and *Caecilia*^{7,13}, and the distribution of fetal and fetal-like teeth across viviparous and oviparous caecilians is consistent with their having a single origin and thus being homologous (Supplementary Information). This implies that the independently derived lineages of viviparous caecilians evolved from (possibly maternal dermatotrophic) ancestors that already possessed a specialized dentition that was preadapted to feeding in oviducts. This reconstruction and current estimates of divergence times²² indicates that fetal-like teeth evolved in the Mesozoic and that some form of skin feeding might have persisted in caecilians for at least 150 million years.

The use of fetal-like teeth in other oviparous caecilians that possess them, whether in maternal dermatotrophy as exemplified by *B. taitanus* or in some other kind of feeding, has not yet been documented. Newborns of the viviparous west African caecilian genus *Geotrypetes* are altricial and it has been speculated, but not shown, that they might feed on the maternal skin or its secretions^{14,23}. A single reported newborn of the viviparous east African caecilian genus *Scolecomorphus* has a peculiar oral morphology that might be associated with specialized feeding after parturition²⁴. Careful observation of these and other as yet unstudied caecilians may reveal additional forms of parental care that are plausible intermediates between, or might otherwise help to explain, the major evolutionary transitions in caecilian reproduction.

One potential advantage to feeding young rather than providing them with yolk alone, is that investment can be delayed and, if advantageous, redirected. Both maternal dermatotrophic and viviparous caecilians produce fewer, larger independent offspring than lecithotrophic caecilians (A.K., unpublished). Selection for larger

offspring is proposed to have driven the evolution of extended parental care in salamanders⁵ and might have similarly driven the evolution of the peculiar derived life histories in caecilians.

In recent years, the known species diversity of amphibians has been steadily increasing, mainly as a result of biodiversity surveys in the tropics²⁵. At the same time there has been growing concern about apparently declining amphibian populations worldwide. Recently the Global Amphibian Assessment identified many data-deficient species (20%) and the urgent need for more information^{26,27}. Our discovery underscores the need for further studies to improve the documentation of the amazing diversity of amphibian life-history strategies and for greater efforts to conserve it.

METHODS

We studied the caecilian *B. taitanus* in the field in southeastern Kenya (Wundanyi, Taita hills, Taita-Taveta District). Most fieldwork was performed after the short rainy season (Vuli), in January 2004 and 2005 after preliminary fieldwork in January 1996. Field-collected females and their young were housed in small plastic boxes (9 cm × 9 cm × 3.5 cm) containing earth moulded to resemble nests observed in the field. Observations were made daily from 06:00 to 09:00 and from 20:00 to 00:00. Behaviours were recorded with a digital video camera (Sony DCR-HC40E). The total length of young during parental care was measured to the nearest millimetre on plastic-coated graph paper. Female body mass was recorded with a digital balance.

Skin tissue of brooding and non-brooding females was fixed in buffered formalin and/or Bouin's fixative. Samples were embedded in accordance with standard procedures²⁸. Paraffin sections (6–8 µm) were cut with a rotary microtome and stained with either haematoxylin/eosin, sudan black B, bromophenol blue, alcian blue pH 2.5 or periodic acid-Schiff.

Tooth morphology of young and adults was examined with a scanning electron microscope (Hitachi 2500 series). Samples were transferred through an acetone series and critical-point dried with carbon dioxide, mounted on aluminium stubs and sputter-coated with gold-palladium.

Received 28 September; accepted 7 November 2005.

- Clutton-Brock, T. H. *The Evolution of Parental Care* (Princeton Univ. Press, Princeton, New Jersey, 1991).
- Smith, C. C. & Fretwell, S. D. The optimal balance between size and number of offspring. *Am. Nat.* **108**, 499–506 (1974).
- Nussbaum, R. A. & Hinkel, H. Revision of East African caecilians of the genera *Afrocaecilia* Taylor and *Boulengerula* Törner (Amphibia: Gymnophiona: Caeciliidae). *Copeia* **1994**, 750–760 (1994).
- Lehtinen, R. M. & Nussbaum, R. A. in *Reproductive Biology and Phylogeny of Anura* (ed. Jamieson, B. G. M.) 343–386 (Science Publishers, Enfield, 2003).
- Nussbaum, R. A. in *Reproductive Biology and Phylogeny of Urodela* (ed. Sever, D. M.) 527–612 (Science Publishers, Enfield, 2003).
- Parker, H. W. Viviparous caecilians and amphibian phylogeny. *Nature* **178**, 250–252 (1956).
- Parker, H. W. & Dunn, E. R. Dentitional metamorphosis in the Amphibia. *Copeia* **1964**, 75–86 (1964).
- Taylor, E. H. *Caecilians of the World* (Kansas Univ. Press, Lawrence, 1968).
- Welsch, U., Müller, M. & Schubert, C. Elektronenmikroskopische und histochemische Beobachtungen zur Fortpflanzungsbiologie viviparer Gymnophionen (*Chthonerpeton indistinctum*). *Zool. Jb. Anat.* **97**, 532–549 (1977).
- Wake, M. H. & Dickie, R. Oviduct structure and function and reproductive modes in amphibians. *J. Exp. Zool.* **282**, 477–506 (1998).
- Exbrayat, J.-M. *Les Gymnophiones, ces curieux Amphibiens* (Edition Boubée, Paris, 2000).
- Himstedt, W. *Die Blindwühlen* (Westarp, Magdeburg, 1996).
- Wilkinson, M. & Nussbaum, R. A. Caecilian viviparity and amniote origins. *J. Nat. Hist.* **32**, 1403–1409 (1998).
- Pennisi, E. Society for Integrative and Comparative Biology: meeting spotlights creatures great and small. *Science* **283**, 623–625 (1999).
- Malonza, P. K. & Measey, G. J. Life history of an African caecilian: *Boulengerula taitanus* Loveridge 1935 (Amphibia Gymnophiona Caeciliidae). *Trop. Zool.* **18**, 49–66 (2005).
- Müller, H., Oommen, O. V. & Bartsch, P. Skeletal development of the direct developing caecilian *Gegeneophis ramsawamii* (Amphibia: Gymnophiona: Caeciliidae). *Zoology* **124**, 171–188 (2005).
- Weldon, P. J., Demeter, B. J. & Rosscoe, R. A survey of shed skin-eating (dermatophagy) in amphibians and reptiles. *J. Herpetol.* **27**, 219–228 (1993).
- Toledo, R. C. & Jared, C. Cutaneous adaptations to water balance in amphibians. *Comp. Biochem. Physiol.* **105**, 593–608 (1993).
- Toledo, R. C. & Jared, C. Cutaneous granular glands and amphibian venoms. *Comp. Biochem. Physiol.* **111**, 1–29 (1995).

20. Kupfer, A., Nabhitabhata, J. & Himstedt, W. Reproductive ecology of female caecilian amphibians (genus *Ichthyophis*): a baseline study. *Biol. J. Linn. Soc.* **83**, 207–217 (2004).
21. Wilkinson, M., Loader, S. P., Gower, D. J., Sheps, J. A. & Cohen, B. L. Phylogenetic relationships of African caecilians (Amphibia: Gymnophiona): insights from mitochondrial rRNA gene sequences. *Afr. J. Herpetol.* **52**, 83–92 (2003).
22. San Mauro, D., Vences, M., Alcobendas, M., Zardoya, R. & Meyer, A. Initial diversification of living amphibians predated the breakup of Pangaea. *Am. Nat.* **65**, 590–599 (2005).
23. O'Reilly, J. C., Fenolio, D., Rania, L. C. & Wilkinson, M. Altriciality and extended parental care in the West African caecilian *Geotrypetes seraphini* (Gymnophiona: Caeciliidae). *Am. Zool.* **38**, 187A (1998).
24. Loader, S. P., Wilkinson, M., Gower, D. J. & Msuya, C. A. A remarkable young *Scolecophorus vittatus* (Amphibia: Gymnophiona: Scolecophoridae) from the North Pare Mountains, Tanzania. *J. Zool. (Lond.)* **259**, 93–101 (2003).
25. Haddad, C. F. B. & Prado, C. P. A. Reproductive modes in frogs and their unexpected diversity in the Atlantic Forest of Brazil. *Bioscience* **55**, 207–217 (2005).
26. Stuart, S. N. *et al.* Status and trends of amphibian declines and extinctions worldwide. *Science* **306**, 1783–1786 (2004).
27. Gower, D. J. & Wilkinson, M. Conservation biology of caecilian amphibians: a review. *Conserv. Biol.* **19**, 45–55 (2005).
28. Romeis, B. *Mikroskopische Technik* (Urban, München, 1989).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Rotich, A. H. Jama and J. Western for arranging collection and export permits; A. Espira, J. Kibirisho and P. W. Kibirisho, J. W. Maghanga, A. Mschimba, D. Mwaghanja and B. Mwakina for access to their land and for help in performing fieldwork; A. Ball, E. B. Morello, J. Newberry and B. Williamson for help with the SEM preparations; D. Cooper for preparing some histological sections; T. Vinhas for help with processing video footage; B. Bwong, P. K. Malonza and G. J. Measey for logistic support; and R. Britz, J. J. Day, D. J. Gower, S. Mohun, L. Rüber and E. Valk for improving earlier versions of the manuscript. This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico, by the Natural Environment Research Council, and by a Marie Curie Fellowship.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.W. (mw@bmn.org).

LETTERS

Independent evolution of bitter-taste sensitivity in humans and chimpanzees

Stephen Wooding¹, Bernd Bufe², Christina Grassi³, Michael T. Howard¹, Anne C. Stone⁴, Maribel Vazquez³, Diane M. Dunn¹, Wolfgang Meyerhof², Robert B. Weiss¹ & Michael J. Bamshad^{1†}

It was reported over 65 years ago that chimpanzees, like humans, vary in taste sensitivity to the bitter compound phenylthiocarbamide (PTC)¹. This was suggested to be the result of a shared balanced polymorphism, defining the first, and now classic, example of the effects of balancing selection in great apes. In humans, variable PTC sensitivity is largely controlled by the segregation of two common alleles at the *TAS2R38* locus, which encode receptor variants with different ligand affinities^{2–4}. Here we show that PTC taste sensitivity in chimpanzees is also controlled by two common alleles of *TAS2R38*; however, neither of these alleles is shared with humans. Instead, a mutation of the initiation codon results in the use of an alternative downstream start codon and production of a truncated receptor variant that fails to respond to PTC *in vitro*. Association testing of PTC sensitivity in a cohort of captive chimpanzees confirmed that chimpanzee *TAS2R38* genotype accurately predicts taster status *in vivo*. Therefore, although Fisher *et al.*'s observations¹ were accurate, their explanation was wrong. Humans and chimpanzees share variable taste sensitivity to bitter compounds mediated by PTC receptor variants, but the molecular basis of this variation has arisen twice, independently, in the two species.

Sensitivity to bitter tastes provides an important means for animals to interact with their environment^{5–7}. By allowing the detection of various toxic compounds in food, especially noxious compounds produced by plants as a means of defence against herbivores, bitter-taste sensitivity enables animals to regulate their intake of toxins^{8,9}. Consequently, bitter-taste perception allows animals to exploit nutritious but toxic food sources by monitoring the consumption of compounds that might cause sickness or death. For humans today, bitter-taste sensitivity is probably less important for avoiding poisoning, but might still influence health through its effects on diet choice⁹ and other behaviours, such as smoking^{10,11}.

In mammals, the ability to taste bitter compounds is mediated principally by a series of small G-protein-coupled receptors encoded by the *TAS2R* gene family that are expressed in specialized taste bud cells in the lingual epithelium⁷. Inter-individual differences in bitter-taste sensitivity arise by a variety of mechanisms, including substantial intra- and inter-genic variation in the *TAS2R* genes^{12–14}. For example, genetic variants that code for functionally distinct receptor types contribute to variation in bitter-taste sensitivity in both humans and mice^{2,4,15,16}. Several *T2R* variants seem to have experienced selective pressures as humans moved into new environments during their dispersal from sub-Saharan Africa about 100,000 years ago^{14,17,18}. Other primate species show variation in their taste sensitivity to bitter compounds as well¹, suggesting that

natural selection might have affected patterns of variation in *T2R* genes even earlier in hominid evolution.

In 1939, Fisher *et al.* reported that chimpanzees and humans both harbour apparently dominant and recessive 'taster' and 'nontaster' alleles at roughly equal frequencies. They argued that the best explanation for this finding is that humans and chimpanzees share anciently derived alleles that evolved before the human–chimpanzee divergence (Fig. 1a) and have since been maintained at equilibrium as the result of heterozygote advantage: "Wherein the selective advantages lie, it would at present be useless to conjecture, but of the existence of a stably balanced and enduring dimorphism determined by this gene there can be no room for doubt."¹

It has recently been reported² that 50–85% of the phenotypic variance in PTC sensitivity in humans is attributable to polymorphisms at the *TAS2R38* locus, in h*TAS2R38*, a ~1-kb single-exon gene on chromosome 7. Most of this variance is explained by amino acid polymorphisms at positions 49 (encoding proline or alanine), 262 (encoding alanine or valine) and 296 (encoding valine or isoleucine) that give rise to two common isoforms, denoted PAV and AVI (haplotypes A and G in Fig. 2a)^{2,3}. h*TAS2R38*(PAV) and h*TAS2R38*(AVI) are strongly associated with PTC sensitivity in human subjects³. *In vitro* stimulation of cells expressing h*TAS2R38*(PAV) with PTC increases cytosolic Ca²⁺ concentrations, whereas stimulation of cells expressing h*TAS2R38*(AVI) does not⁴. To test the hypothesis that PTC sensitivity in chimpanzees is controlled by the

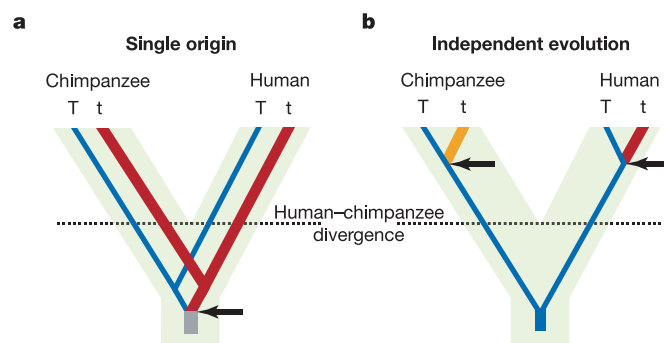


Figure 1 | Hypothetical origins of PTC taster and nontaster alleles in humans and chimpanzees. **a**, Divergence of taster (T) and nontaster (t) alleles before the human–chimpanzee divergence, with T and t alleles subsequently maintained within each species by balancing natural selection. **b**, Independent origin of t alleles from T alleles in humans and chimpanzees, after the human–chimpanzee divergence.

¹Department of Human Genetics, University of Utah, 15 North 2030 East, Salt Lake City, Utah 84112-5330, USA. ²German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, Nuthetal 14558, Germany. ³Department of Comparative Medicine, Southwest Foundation for Biomedical Research, San Antonio, Texas 78245-0549, USA. ⁴School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona 85287-2402, USA. †Present address: Departments of Pediatrics and Genome Sciences, University of Washington, Seattle, Washington 98195, USA.

same taster and nontaster alleles found in humans¹, we re-sequenced the hTAS2R38 orthologue in chimpanzees¹⁹ (chTAS2R38) in 37 wild-born animals representing three subspecies (26 *Pan troglodytes verus*, 9 *P. t. troglodytes* and 2 *P. t. schweinfurthii*) and 49 captive-born animals (*P. t. verus*), and compared the molecular and functional characteristics of human and chimpanzee alleles.

We found seven variable nucleotide positions in chTAS2R38 (Fig. 2a). These did not occur at any of the nucleotide or amino acid positions variable in humans, including the positions distinguishing the hTAS2R38(PAV) and hTAS2R38(AVI) isoforms. However, one site—a T→G substitution found at a frequency of 0.52 in the second position of the initiation codon (that is, ATG→AGG)—is noteworthy because it is predicted to attenuate

chTAS2R38 translation. On the basis of this finding, we hypothesized that the chTAS2R38 haplotypes harbouring the AGG variant were chimpanzee PTC nontaster alleles (Fig. 2a).

A maximum-likelihood phylogeny revealed that TAS2R38 haplotypes in humans and chimpanzees form separate clades, and that the putative nontaster allele within each species is evolutionarily derived from an ancestral taster allele (Fig. 2b). In addition, the application of Tajima's *D* test to the data from wild-born chimpanzees failed to reject the hypothesis of evolutionary neutrality ($D = -0.59$, $P < 0.40$), yielding a *D* value well within the range of *D* values reported for 59 noncoding regions thought to be evolving neutrally in chimpanzees (-0.39 to -1.94 ; refs 20–22). This result stands in contrast to findings for TAS2R38 in humans, for which the value of Tajima's *D* is 1.55. This value is (1) significantly greater than expected ($P < 0.01$), (2) greater than 47 of 50 *D* values in noncoding regions, and (3) greater than 99.5% of *D* values previously reported for more than 1,600 genes²³. The high value of *D* in humans has been interpreted as evidence for balancing selection acting on hTAS2R38 (ref. 17). These findings suggest that not only have nontaster alleles been independently derived in humans and chimpanzees, but they seem to have arisen under different selective regimes. Thus, we reject Fisher *et al.*'s hypothesis that humans and chimpanzees share anciently derived taster and nontaster alleles maintained by balancing natural selection.

To further explore the evolutionary implications of the ATG→AGG mutation in chTAS2R38, we investigated its functional consequences for translation initiation and receptor activity. *In vitro* transcription and translation of a chTAS2R38–reporter fusion construct confirmed that a full-length TAS2R38 protein (chTAS2R38_{WT}–R) was produced by the wild-type chTAS2R38(ATG) allele, but not by the chTAS2R38(AGG) allele (Fig. 3a and Supplementary Fig. S1). Instead, the main product of the chTAS2R38(AGG) allele was a truncated polypeptide, chTAS2R38_{TR}–R, resulting from translation initiation at a downstream ATG codon (M97), the expression of which could be abolished by substituting a leucine (CTG) codon at that position (Fig. 3a). chTAS2R38_{TR}–R is predicted to lack 96 amino-terminal residues of the chTAS2R38_{WT}–R receptor, suggesting that the ATG→AGG mutation severely reduces or abolishes the production of functional PTC receptor in chimpanzees.

To test the functional properties of the chTAS2R38–R receptor variants, we cloned the ATG and AGG alleles of chTAS2R38 from the genomic DNA of two homozygous chimpanzees as fusion proteins with an N-terminal SST-tag (which facilitates cell-surface expression) and a C-terminal Flag tag to allow immunological detection of the recombinantly expressed fusion proteins²⁴ (Supplementary Fig. S2). These constructs were then transiently transfected into HEK-293T cells that stably express the chimaeric G-protein subunit $G_{\alpha 16\text{gust44}}$, which has been shown to efficiently couple TAS2R receptors to the release of intracellular calcium²⁵. Immunohistochemical analysis revealed that all receptor variants were expressed, and no obvious differences in cell-surface expression levels were detected (Fig. 3b and Supplementary Figs S3, S4).

Fluorescence imaging plate reader recordings of calcium transients in cells expressing chTAS2R38 receptor variants showed that micromolar concentrations of PTC increased cytosolic Ca^{2+} in cells transfected with chTAS2R38(ATG) in a concentration-dependent manner. These analyses showed that although the dose–response curve of the predicted chimpanzee taster allele, chTAS2R38(ATG), was similar to that of the human taster allele, hTAS2R38(PAV), cells expressing chTAS2R38(AGG) failed to respond to PTC even at concentrations as high as 100 μM (Fig. 3c and Supplementary Fig. S4). Engineering an initiation codon upstream of the native start site in the chTAS2R38(AGG) allele resulted in complete rescue of receptor function (Fig. 3b, c). These results confirm that chTAS2R38(ATG) produces a functional PTC receptor, but the chTAS2R38(AGG) allele does not.

To determine whether chTAS2R38(AGG) is the nontaster allele

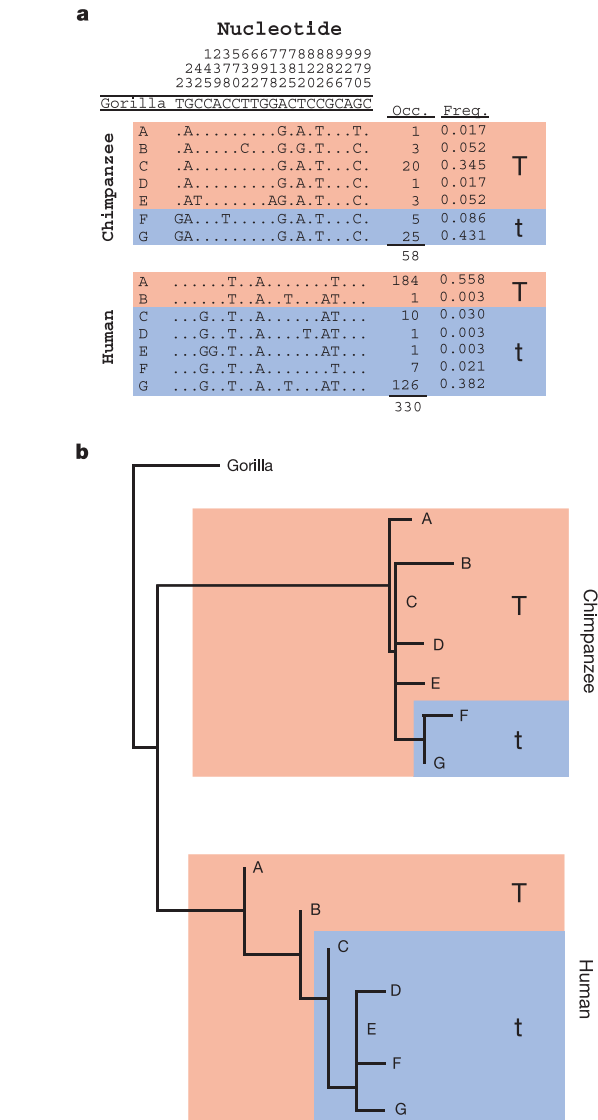


Figure 2 | Patterns of genetic variation in TAS2R38 in human, chimpanzee and gorilla. **a**, The nucleotide sequence of each TAS2R38 haplotype compared to that of gorilla (top). Dots indicate sequence identity and letters indicate differences. Human and chimpanzee TAS2R38 haplotypes differed by a minimum of six nucleotides. No nucleotide position was variable in both humans and chimpanzees. T and t denote 'taster' and 'nontaster' alleles, respectively. **b**, Maximum-likelihood phylogeny of human and chimpanzee haplotypes generated using PHYLIP and using the gorilla as an outgroup. Linkage disequilibrium was high ($D' \sim 1.0$) across the length of the gene (~ 1 kb). The human and chimpanzee TAS2R38 alleles form two separate clades.

in vivo, we tested whether *chTAS2R38* genotype predicted PTC sensitivity in chimpanzees. PTC sensitivity was measured in 42 unrelated captive-born chimpanzees by presenting each animal with apple slices soaked in control and test solutions (water and 4.0 mM PTC in water, respectively). Apples were chosen because they are a typical enrichment food for chimpanzees and are easily

saturated with PTC. The response of each chimpanzee was then scored on a categorical scale from 1 (ready acceptance) to 5 (strong rejection) by a primate behavioural specialist who directs the enrichment program for these chimpanzees, has day-to-day contact with the animals and is familiar with their dietary habits (see Methods for details). The scorer was blinded to chimpanzee genotype.

Among the 39 chimpanzees for which complete genotype and phenotype data were available, 11 had an ATG/ATG genotype, 19 had ATG/AGG and 9 had AGG/AGG (Supplementary Notes S1 and S2). Preliminary association analyses comparing only the ATG/ATG and ATG/AGG animals showed that these genotypes responded similarly to the PTC treatment, suggesting that ATG is dominant or nearly dominant to AGG (Supplementary Note S3). This observation is consistent with the results of our *in vitro* translation experiments, which suggest that *chTAS2R38*(AGG) is effectively a null allele. For this reason, we grouped the ATG/ATG and ATG/AGG animals as predicted tasters in all subsequent analyses, and AGG/AGG animals were predicted nontasters.

Of the 30 predicted tasters, 18 showed the taster phenotype (that is, responses ranging from 3 to 5 on our scale) (Fig. 4). Of the 9 chimpanzees predicted to be PTC nontasters, all showed the nontaster phenotype (that is, responses of 1 or 2 on our scale) (Fig. 4 and Supplementary Note S4). A one-sided Fisher's exact test revealed a significant association between *chTAS2R38* genotype and response to the test (PTC) solution ($P < 0.015$), but not response to the control (water) solution ($P > 0.75$). When the experiment was repeated, 23 out of 30 predicted tasters showed the taster phenotype, and 8 of 9 predicted nontasters showed the nontaster phenotype (Fig. 4). These results confirm a significant association between genotype and response to the PTC solution ($P < 0.01$) but not the control solution ($P > 0.75$).

A number of predicted tasters did not reject PTC-treated apples in one or both tests. We hypothesize that these animals either have PTC taste thresholds higher than that tested in this study (4.0 mM, which is less than half of the maximum concentration of 8.5 mM previously tested in humans^{2,3}), or that they could taste the compound but were not noticeably disturbed by it. The assignment of these predicted tasters to the nontaster outcome group makes our association tests statistically more conservative. However, to verify the association test, we compared the responses of predicted tasters and nontasters separately. These tests showed that although predicted tasters showed a strong difference in response between the test and control

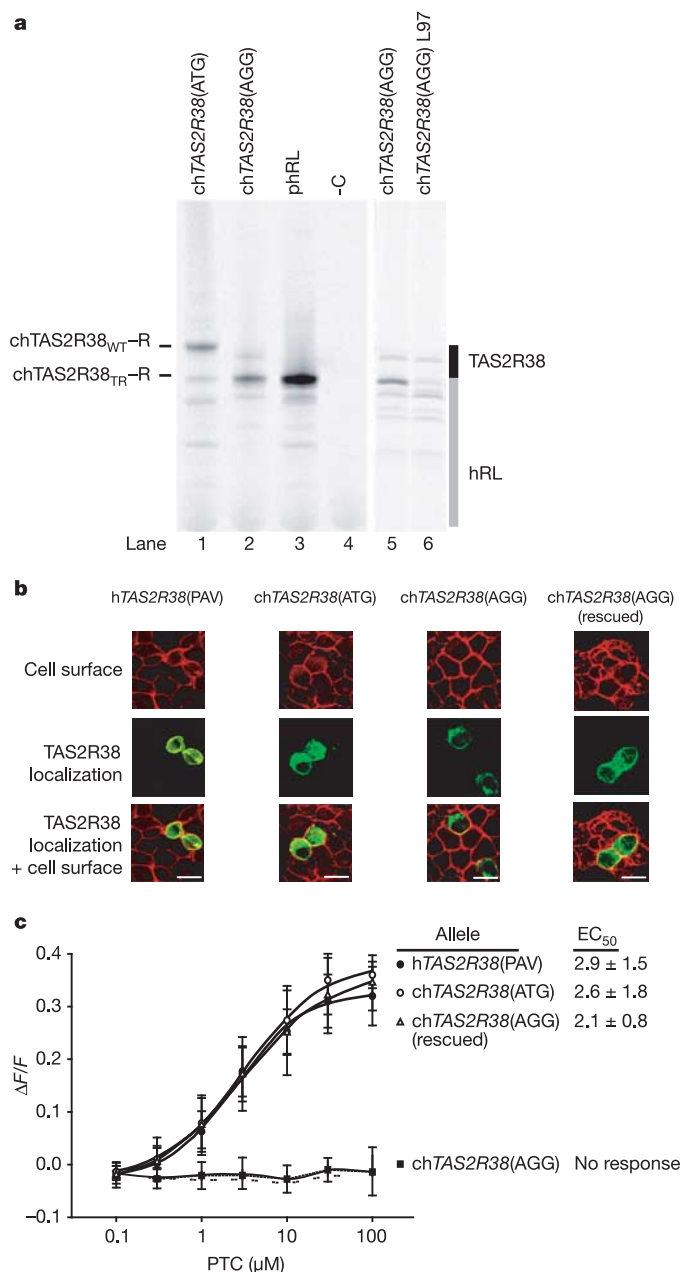


Figure 3 | Translation, membrane localization and functional characteristics of *chTAS2R38* alleles. **a**, Lane 1, full-length product (*chTAS2R38*_{WT}) of *chTAS2R38*(ATG). Lanes 2 and 5, truncated product (*chTAS2R38*_{TR}) of *chTAS2R38*(AGG). Lane 3, vector. Lane 4, negative control. Lane 6, product of *chTAS2R38*(AGG) with ATG at position 97 changed to CTG. Predicted products of TAS2R38 (black bar) and *Renilla* luciferase (grey bar). **b**, Confocal fluorescence images of HEK-293T cells transfected with TAS2R38 allele. Cell-surface staining is red, PTC receptor is green, overlay of PTC receptor on the cell surface is yellow. Scale bar, 10 μm. **c**, Dose-response curves of increasing PTC concentration on intracellular Ca²⁺ levels (measured as change in fluorescence, $\Delta F/F$) in cells expressing TAS2R38. Each point represents mean $\pm$ s.e.m. of three independent replicates.

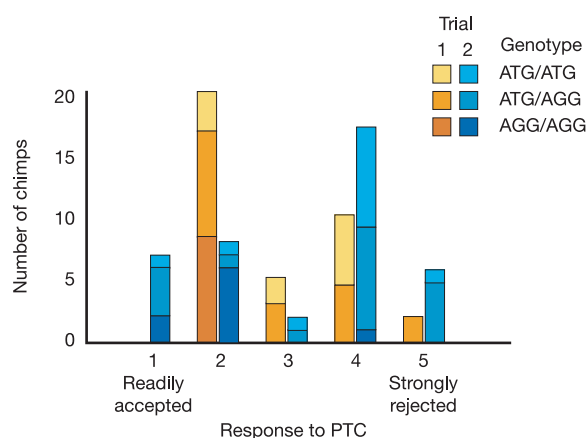


Figure 4 | Chimpanzee responses to PTC. Distribution of scored responses to apple slices soaked in 4.0 mM PTC solution, by genotype. Responses in tests 1 and 2 are shown in yellow and blue columns, respectively. Genotype is indicated by shading: light, ATG/ATG; intermediate, ATG/AGG; dark, AGG/AGG.

treatments (Fisher's exact test; $P < 10^{-5}$ for both the first and second trials), nontasters showed no difference ($P > 0.45$ for both the first and second trials) (Supplementary Note S5). Thus, *TAS2R38* genotype is highly predictive of PTC taste sensitivity in chimpanzees.

On the basis of our results, we reject Fisher *et al.*'s long-standing hypothesis that the variation in PTC taste sensitivity shared between humans and chimpanzees is due to shared ancient genetic polymorphisms maintained by balancing natural selection. Phylogenetic relationships among chimpanzee and human haplotypes are more consistent with the hypothesis that nontaster alleles have evolved at least twice, independently, over the course of hominid evolution (Fig. 1b). Whether this finding is specific to *TAS2R38* or is a more generalized trend among bitter taste receptors remains to be determined. Characterizing such patterns could facilitate the discovery of compounds with particularly important effects on nutrition and health. Access to the recently sequenced chimpanzee genome and a captive cohort of healthy chimpanzees offers the opportunity to explore this issue specifically, and more broadly paves the way for future comparative functional genomic studies with our sibling species.

METHODS

DNA sequence analysis. Haplotypes were inferred from unphased genotypes in wild-born chimpanzees using the PHASE 2.1 computer program²⁶. Tajima's *D* test of evolutionary neutrality²⁷ was performed using the Dfsc computer program¹⁷. This test was performed under the assumption that the effective population size of *P. t. verus*, which represents the majority of the wild-born sample, has remained constant. This assumption is based on several studies suggesting that population sizes in western, but not eastern, chimpanzee subspecies have remained stable for an extended period^{20,28–30}. Tajima's *D* value at *chTAS2R38* was also compared with values calculated for 50 noncoding regions, each more than 5 kb from any known coding region, which were assumed to be evolving neutrally²².

Expression studies. The last 88 nucleotides of the 5' untranslated region and first 318 coding nucleotides of *chTAS2R38* alleles were subcloned into the *NheI* and *AvaI* restriction sites of the pRL-CMV plasmid vector (Promega). These constructs and the parent vector (pRL) were transcribed and translated *in vitro* using a coupled transcription and translation system (T7 TNT, Promega) in the presence of ³⁵S-methionine. ³⁵S-labelled protein products were separated by SDS polyacrylamide gel electrophoresis.

Functional assay. The allele encoding *chTAS2R38*_{TR-R} was amplified from genomic DNA of a chimpanzee homozygous for the *chTAS2R38*(AGG) (nontaster) allele, and the *chTAS2R38*_{WT-R} receptor was amplified from a *chTAS2R38*(ATG) homozygote using standard polymerase chain reaction (PCR) conditions and primer sets directed against the coding region of the respective receptor variants. A primer-encoded *EcoRI* site in front of the start codon and a primer-encoded *NotI* site in front of the stop codon were used to subclone the PCR products into a pCDNA5 expression vector containing an N-terminal SST and a C-terminal Flag tag (Supplementary Fig. 2). All cloned receptors were sequenced to exclude amplification errors during the PCR reaction.

Functional assays were carried out essentially as described in ref. 18. Plasmids containing the receptor variants were transiently transfected into HEK-293T cells stably expressing the chimaeric G-protein subunit *G_{α16gust44}* using Lipofectamine 2000 (Invitrogen). Calcium mobilization of Fluo-4-AM loaded cells upon stimulation was monitored with an automated fluorometric imaging plate reader (FLIPR, Molecular Devices). The obtained calcium signals were corrected for background fluorescence and the response of mock-transfected cells. Concentration–response curves and *EC*₅₀ values were averaged over three independent experiments and were calculated in Sigma Plot by nonlinear regression using the function $f(x) = [(a - d)/(1 + (x/EC_{50})^{nh}) + d]$.

Immunocytochemistry. Immunocytochemistry was carried out as described in ref. 18. HEK-293T cells expressing *G_{α16gust44}* were seeded on coverslips transfected with plasmids containing the different *TAS2R38* receptor variants. Twenty-four hours later, the cells were permeabilized using a 1:1 acetone:methanol solution. The receptor variants were detected with a primary monoclonal mouse anti Flag-antibody (1:1,000; Sigma) and a secondary Alexa-488-conjugated goat antiserum against mouse IgG (1:1,000; Molecular Probes). Expression levels were determined on three independent days, analysing three independent visual fields per variant using a fluorescence microscope (Zeiss Axioplan) and a cooled CCD camera (Visitron Systems). To assess cell-surface expression, cells were additionally incubated on ice with 5 μg ml⁻¹ biotin-

labelled concanavalin A (Sigma) for 1 h before permeabilization. The cell surface was visualized with streptavidin-Alexa-633 (1:1,000; Molecular Probes) (Fig. 3b). Colocalization between the receptor and cell surface was analysed using a Leica TCS SP2 confocal microscope.

Phenotyping. Apples were cored, separated into 14 slices of equal proportion, and soaked in either 2 l of water (control solution) or 2 l of a 4.0 mM PTC solution for 10–12 h overnight before testing. Solutions were prepared with different utensils and in different work areas to minimize cross-contamination. On each test day, the excess solution was drained and individual apple slices were offered to chimpanzee test subjects. This procedure recapitulated the manner in which enrichment foods (for example, grapes and apples) were typically provided to the chimpanzees. Each chimpanzee was monitored by a trained observer until either the apple slice was consumed or rejected, for a maximum period of ~10 min. The response of each chimpanzee was then scored on a categorical scale from 1 (ready acceptance) to 5 (strong rejection) based on whether the apple slice was consumed, the time taken to consume the apple slice, the presence of excessive salivation, and the extent of facial grimacing. Testing of the control and test solutions was alternately performed on different days, each separated by breaks of ~48 h. Immediately after testing with either solution, a second enrichment food (dried apricot without PTC) was provided to each test subject.

Received 17 November 2005; accepted 16 February 2006.

1. Fisher, R. A., Ford, E. B. & Huxley, J. Taste-testing the anthropoid apes. *Nature* **144**, 750 (1939).
2. Kim, U. K. *et al.* Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science* **299**, 1221–1225 (2003).
3. Drayna, D. *et al.* Genetic analysis of a complex trait in the Utah Genetic Reference Project: a major locus for PTC taste ability on chromosome 7q and a secondary locus on chromosome 16p. *Hum. Genet.* **112**, 567–572 (2003).
4. Bufe, B. *et al.* The molecular basis of individual differences in phenylthiocarbamide and propylthiouracil bitterness perception. *Curr. Biol.* **15**, 322–327 (2005).
5. Meunier, N., Marion-Poll, F., Rospars, J. P. & Tanimura, T. Peripheral coding of bitter taste in *Drosophila*. *J. Neurobiol.* **56**, 139–152 (2003).
6. Hilliard, M. A., Bergamasco, C., Arbucci, S., Plasterk, R. H. & Bazzicalupo, P. Worms taste bitter: ASH neurons, QUL-1, GPA-3 and ODR-3 mediate quinine avoidance in *Caenorhabditis elegans*. *EMBO J.* **23**, 1101–1111 (2004).
7. Hoon, M. A. *et al.* Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity. *Cell* **96**, 541–551 (1999).
8. Gendinning, J. I. Is the bitter rejection response always adaptive? *Physiol. Behav.* **56**, 1217–1227 (1994).
9. Drewnowski, A. & Rock, C. L. The influence of genetic taste markers on food acceptance. *Am. J. Clin. Nutr.* **62**, 506–511 (1995).
10. Kaplan, A. R., Glanville, E. V. & Fischer, R. Taste thresholds for bitterness and cigarette smoking. *Nature* **202**, 1366 (1964).
11. Enoch, M. A., Harris, C. R. & Goldman, D. Does a reduced sensitivity to bitter taste increase the risk of becoming nicotine addicted? *Addict. Behav.* **26**, 399–404 (2001).
12. Conte, C., Ebeling, M., Marcuz, A., Nef, P. & Andres-Barquin, P. J. Evolutionary relationships of the *Tas2r* receptor gene families in mouse and human. *Physiol. Genomics* **14**, 73–82 (2003).
13. Parry, C. M., Erkner, A. & le Coutre, J. Divergence of *T2R* chemosensory receptor families in humans, bonobos, and chimpanzees. *Proc. Natl Acad. Sci. USA* **101**, 14830–14834 (2004).
14. Kim, U., Wooding, S., Ricci, D., Jorde, L. B. & Drayna, D. Worldwide haplotype diversity and coding sequence variation at human bitter taste receptor loci. *Hum. Mutat.* **26**, 199–204 (2005).
15. Adler, E. *et al.* A novel family of mammalian taste receptors. *Cell* **100**, 693–702 (2000).
16. Chandrashekar, J. *et al.* *T2Rs* function as bitter taste receptors. *Cell* **100**, 703–711 (2000).
17. Wooding, S. *et al.* Natural selection and molecular evolution in *PTC*, a bitter-taste receptor gene. *Am. J. Hum. Genet.* **74**, 637–646 (2004).
18. Soranzo, N. *et al.* Positive selection on a high-sensitivity allele of the human bitter-taste receptor *TAS2R16*. *Curr. Biol.* **15**, 1257–1265 (2005).
19. Wang, X., Thomas, S. D. & Zhang, J. Relaxation of selective constraint and loss of function in the evolution of human bitter taste receptor genes. *Hum. Mol. Genet.* **13**, 2671–2678 (2004).
20. Fischer, A., Wiebe, V., Paabo, S. & Przeworski, M. Evidence for a complex demographic history of chimpanzees. *Mol. Biol. Evol.* **21**, 799–808 (2004).
21. Gilad, Y., Bustamante, C. D., Lancet, D. & Paabo, S. Natural selection on the olfactory receptor gene family in humans and chimpanzees. *Am. J. Hum. Genet.* **73**, 489–501 (2003).
22. Yu, N. *et al.* Low nucleotide diversity in chimpanzees and bonobos. *Genetics* **164**, 1511–1518 (2003).
23. Schneider, J. A. *et al.* DNA variability of human genes. *Mech. Ageing Dev.* **124**, 17–25 (2003).

24. Bufe, B., Hofmann, T., Krautwurst, D., Raguse, J. D. & Meyerhof, W. The human TAS2R16 receptor mediates bitter taste in response to β -glucopyranosides. *Nature Genet.* **32**, 397–401 (2002).
25. Ueda, T., Ugawa, S., Yamamura, H., Imaizumi, Y. & Shimada, S. Functional interaction between T2R taste receptors and G-protein α subunits expressed in taste receptor cells. *J. Neurosci.* **23**, 7376–7380 (2003).
26. Stephens, M., Smith, N. J. & Donnelly, P. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* **68**, 978–989 (2001).
27. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
28. Gagneux, P. *et al.* Mitochondrial sequences show diverse evolutionary histories of African hominoids. *Proc. Natl Acad. Sci. USA* **96**, 5077–5082 (1999).
29. Goldberg, T. L. & Ruvolo, M. The geographic apportionment of mitochondrial genetic diversity in east African chimpanzees, *Pan troglodytes schweinfurthii*. *Mol. Biol. Evol.* **14**, 976–984 (1997).
30. Kaessmann, H., Wiebe, V., Weiss, G. & Paabo, S. Great ape DNA sequences reveal a reduced diversity and an expansion in humans. *Nature Genet.* **27**, 155–156 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Chimpanzee samples were contributed by Sunset Zoo, Riverside Zoo, Kitwe Point Sanctuary, the Jane Goodall Institute, J. Wickings, the Centre International de Recherches Medicales, the Primate Foundation of Arizona, the New Iberia Primate Center and the Southwest Foundation for Biomedical Research. Chimpanzee samples were imported under a CITES permit. Comments and technical assistance were provided by C. Anderson, D. Drayna, L. Jorde, U.-k. Kim, M. Pyrski, A. Rogers, J. Seger and E. Wooding. Funding was provided by grants from the NIH, the Centers for Disease Control, the Muscular Dystrophy Association, the Wenner-Gren Foundation, the National Science Foundation and the German Science Foundation.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.W. (swooding@genetics.utah.edu) or M.J.B. (mbamshad@u.washington.edu).

STED microscopy reveals that synaptotagmin remains clustered after synaptic vesicle exocytosis

Katrin I. Willig^{1*}, Silvio O. Rizzoli^{2*}, Volker Westphal¹, Reinhard Jahn² & Stefan W. Hell¹

Synaptic transmission is mediated by neurotransmitters that are stored in synaptic vesicles and released by exocytosis upon activation. The vesicle membrane is then retrieved by endocytosis, and synaptic vesicles are regenerated and re-filled with neurotransmitter¹. Although many aspects of vesicle recycling are understood, the fate of the vesicles after fusion is still unclear. Do their components diffuse on the plasma membrane, or do they remain together? This question has been difficult to answer because synaptic vesicles are too small (~40 nm in diameter) and too densely packed to be resolved by available fluorescence microscopes. Here we use stimulated emission depletion (STED)² to reduce the focal spot area by about an order of magnitude below the diffraction limit, thereby resolving individual vesicles in the synapse. We show that synaptotagmin I, a protein resident in the vesicle membrane, remains clustered in isolated patches on the presynaptic membrane regardless of whether the nerve terminals are mildly active or intensely stimulated. This suggests that at least some vesicle constituents remain together during recycling. Our study also demonstrates that questions involving cellular structures with dimensions of a few tens of nanometres can be resolved with conventional far-field optics and visible light.

Synaptic vesicle recycling has been studied for over three decades^{3,4}. In the best documented model of neurotransmitter release, the synaptic vesicle membrane undergoes exocytosis by collapsing into the plasma membrane⁴, referred to as full fusion. Fusion is thought to involve, at least to some extent, mixing and lateral diffusion of vesicle proteins across the plasma membrane before they are re-internalized by endocytosis. Accordingly, to regenerate vesicles, sorting of these proteins must occur at the plasma membrane and/or in an endosomal intermediate^{1,5,6}. However, if vesicle constituents remain patched together after exocytosis, the recycling machinery would just have to internalize the fused vesicular patch without elaborate sorting. To date, it has not been possible to differentiate between these alternatives. Owing to their diffraction-limited resolution, confocal and epifluorescence microscopes cannot identify single-vesicle-derived protein patches in a synaptic bouton; moreover, individual vesicles are also impossible to image, except under sparse labelling conditions^{7–9}. Electron microscopy, on the other hand, provides sufficient resolution, but the required labelling efficiency has hitherto not been achievable.

Recent advances in optical physics have shown that the diffraction barrier of far-field fluorescence microscopy can be broken by stimulated emission depletion (STED)^{2,10,11}. In fact, STED is just one example of a family of microscopy concepts that, in spite of using regular lenses, allow diffraction-unlimited resolution^{11–13}. In a typical STED microscope the excitation beam is overlapped with a doughnut-shaped beam that is capable of de-exciting fluorophores by stimulated emission. Co-alignment of the beams ensures that fluorescence is allowed only in the central area of the excitation

spot where the doughnut beam is close to zero (Fig. 1a). Scanning with a narrowed spot across the sample readily yields subdiffraction images. With a sufficiently intense doughnut, the fluorescent spot of a STED microscope can, in principle, be sharpened down to the molecular scale^{2,11,12}.

To investigate synaptic vesicle recycling in cultured neurons, we built a STED microscope with an objective lens of 1.4 numerical aperture. The green-emitting dye used for antibody labelling was excited at 470 nm and de-excited at 615 nm. Without STED, the full-width-half-maximum (FWHM) of the fluorescent focal spot was 195 nm, which is typical for a first-rate confocal microscope¹⁴. Applying the doughnut of our setup yielded a spot FWHM of 66 nm (Fig. 1a), corresponding to a ninefold reduction in the effective focal area. Moreover, the spot of the STED microscope becomes increasingly thinner towards the top¹², implying that the resolution is slightly better. Simulations revealed that given the encountered noise conditions, our STED microscope allows the separation of point objects that are 45 nm apart in the focal plane, which is sufficient for resolving individual synaptic vesicles within nerve terminals.

For imaging, vesicles were targeted with a monoclonal antibody directed against the intravesicular domain of the synaptic vesicle protein synaptotagmin¹⁵ (Fig. 1b). Figure 1c shows an epifluorescence overview image of primary cultured hippocampal neurons incubated with the antibody at 37 °C for 5 min. Upon neuronal activity, a vesicle opens to the outside space (exocytosis), rendering its inside accessible to the antibodies. These bind to the synaptotagmin molecules and are internalized when the vesicles are retrieved¹⁶. Thus, only vesicles undergoing exocytosis during the incubation time are labelled. These vesicles were visualized by fluorescently labelled secondary antibodies applied after membrane permeabilization and fixation. The gain in resolution becomes apparent by comparing the panels in Fig. 1d showing confocal and STED microscopy images. Whereas the former does not show substructures below the level of individual boutons, the latter resolves numerous dots within the terminals.

To differentiate between the synaptotagmin pool remaining on the surface of the plasma membrane and that internalized by endocytosis, we used two different protocols. First, labelling was performed on ice and in the absence of Ca²⁺; that is, conditions under which no endocytosis occurs. Bright staining was observed (Fig. 2a). Second, labelling was performed at 37 °C to allow for uptake, resulting in antibody binding to both surface-exposed and internalized synaptotagmin pools. To selectively label the internalized pool, we blocked the surface-bound antibodies with unlabelled secondary antibodies before adding dye-labelled secondary antibodies. As we expected, the staining intensity of blocked, non-permeabilized preparations was hardly above background (Fig. 2b). Conversely, strong labelling was observed after permeabilization,

¹Departments of NanoBiophotonics and ²Neurobiology, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany.

*These authors contributed equally to this work.

showing that the internalized (unblocked) vesicles were accessible to secondary antibody labelling (Fig. 2c). To further ensure that internal epitopes were only accessible after permeabilization, we labelled neurons with an antibody specific for the cytoplasmic tail of synaptotagmin¹⁷. Labelling was only observed when cultures were permeabilized before antibody incubation (Fig. 2e, f).

A comparison of surface-exposed and internalized synaptotagmin pools as resolved by STED microscopy is shown in Fig. 2g, h. Both the internalized and the surface-exposed pools were resolved as confined dots. Thus, synaptotagmin remains concentrated in small clusters after exocytosis instead of being dispersed across the plasma membrane. Furthermore, there were consistently fewer dots detectable on the surface compared with the internalized pool, indicating that the plasma membrane pool is short-lived owing to rapid endocytosis. Notably, the dots on the cell surface appeared brighter than those of

the internalized pool. We therefore quantified dot brightness (see Methods). For comparison, a dilute solution of primary antibodies was adsorbed on glass, labelled with secondary antibodies, imaged by STED, and quantified in parallel. The resulting histogram (Fig. 2i) shows (1) that both internalized and surface-exposed dots are significantly brighter than individual antibodies, confirming that each dot represents multiple synaptotagmin molecules, and (2) that the brightness of the surface-exposed patches is shifted towards higher values compared to the internalized vesicles. The second observation is probably due to the brevity of the endocytic process (2–5 s; ref. 18), which limits epitope accessibility for the antibodies. Surface patches, on the other hand, were exposed to the antibody solution for a few minutes (on ice, to inhibit active recycling), resulting in a higher labelling efficiency. Comparison of the internalized pool with the total (unblocked) pool after permeabilization

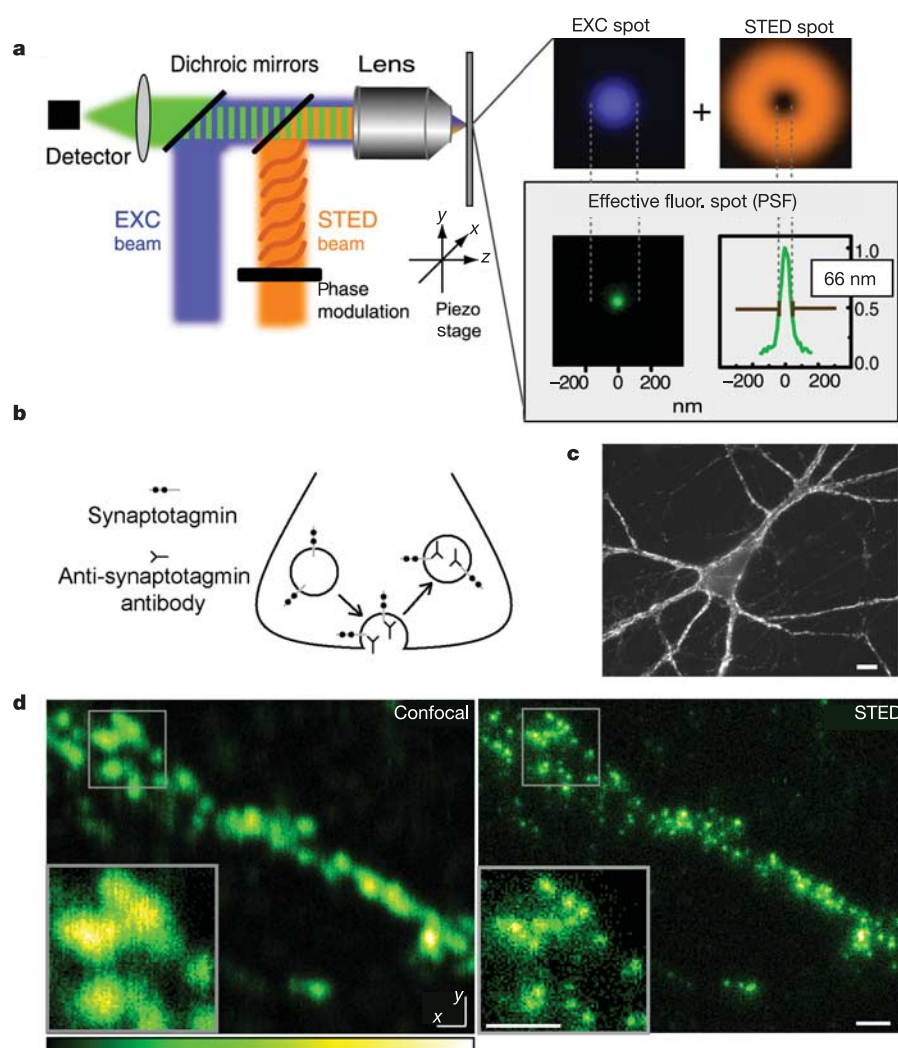


Figure 1 | STED microscopy resolves synaptic vesicles in individual boutons of primary cultured hippocampal neurons. **a**, Principles of operation. While the blue excitation (EXC) beam is focused to a diffraction-limited excitation spot, shown in the adjacent panel in blue, the orange STED beam is able to de-excite molecules. The STED beam is phase-modulated to form the focal doughnut shown in the top right panel. Superimposition of the two focal spots confines the area in which fluorescence is possible to the doughnut centre, yielding the effective fluorescent spot of subdiffraction size shown in green in the lower panel. All spots represent measured data and are drawn to scale. The profile of the green effective fluorescent spot has an FWHM of 66 nm as well as a sharp peak. The green spot shows an 11-fold

reduction in focal area beyond the excitation diffraction value (compare with blue spot). **b**, Mechanism of synaptic labelling. Synaptic vesicles exocytose, allowing their luminal synaptotagmin domains to bind anti-synaptotagmin antibodies. These antibodies are internalized upon endocytosis. **c**, Typical image of a neuron labelled with an anti-synaptotagmin antibody, fixed, permeabilized and visualized using Atto 532-labelled secondary antibodies. Fluorescent puncta represent labelled synaptic nerve terminals. Scale bar, 10 μ m. **d**, Comparison of confocal (left) and STED (right) counterpart images of a labelled preparation reveals a marked increase in resolution by STED. Scale bar, 500 nm.

revealed that a significant fraction of brighter (surface?) dots is preserved, suggesting that permeabilization does not disrupt the surface clusters (Supplementary Fig. S1).

We next strongly stimulated our preparations (using 70 mM KCl) in the presence of anti-synaptotagmin antibodies. Permeabilization of the preparations revealed intense staining, suggesting that many vesicles had been labelled. Total surface fluorescence also increased more than twofold (not shown). However, the staining pattern of the surface pools was very similar to the unstimulated cultures (Fig. 3a, b), indicating that synaptotagmin remains clustered even during high synaptic activity.

We then asked whether the surface dots represent the synaptotagmin inventory of individual vesicles, or whether they represent (at least in part) aggregated patches. As brightness alone cannot be used as a criterion, we compared the dot diameters of the internalized pool with those of the surface pool (Fig. 3c and Supplementary Fig. S2). All populations peaked at 70–85 nm, with no significant differences among them. Considering the 35–40 nm vesicle diameter observed with electron microscopy¹⁹ and the spot size of our STED microscope, we anticipated a dot diameter of ~70 nm. The minority of slightly larger diameters might be due to the bound antibody sandwich. Thus, the observed values are in agreement with values predicted for a membrane patch derived from a single vesicle (Supplementary Fig. S3).

Although our data strongly suggest that synaptotagmin remains clustered after exocytosis, we cannot rule out that the proteins disperse briefly before being re-clustered. However, we regard this option as unlikely because there are only a few spots on the surface that correspond in intensity to single molecules. Previous observations have suggested the dispersal of a vesicle protein (green fluorescent protein (GFP)-tagged synaptobrevin) out of nerve terminals into the axons upon strong stimulation²⁰. However, it remained uncertain whether the dispersal was due to diffusion of patches or of single

molecules. If we further incubate surface-labelled preparations to allow for endocytosis before adding the membrane-impermeant detection reagent, staining is strongly reduced (see Supplementary Fig. S4). Thus, the patches are ultimately endocytosed, possibly through interaction with synaptotagmin²¹.

At present, we do not know whether the behaviour of synaptotagmin is a paradigm for all vesicle proteins. However, vesicle proteins have been reported to adhere to one another after solubilization in various detergents²². Thus, it is conceivable that other vesicle constituents remain associated with synaptotagmin in the plasma membrane when recycling, both at rest and during high neuronal activity. Our findings may also explain the quantal character of vesicle endocytosis described in ref. 7—if the synaptic vesicle components remain clustered in the membrane, it is easy to envisage how endocytosis (and not only exocytosis) can assume quantal behaviour.

Our application of subdiffraction far-field optical resolution to an unsolved problem in cell biology holds great promise for addressing other questions involving dimensions on the scale of tens of nanometres. Maintaining both the labelling efficiency of fluorescence microscopy and its ease of operation, STED should provide an alternative to electron microscopy. Importantly, STED is a purely physical phenomenon; no mathematical data processing is required for the image generation. We also note that the 45–66 nm resolution reported here does not represent a limit in biological imaging. Being inherently diffraction-unlimited, the resolution of a STED microscope can be further increased by optimizing STED and/or the doughnut. In fact, a focal spot width of 16 nm has recently been measured in single-molecule experiments²³. Therefore, our study not only provides insights into the mechanisms of vesicle recycling, but also demonstrates the power of an emerging family of microscopy techniques that, despite using regular objective lenses and visible light, is no longer limited by diffraction.

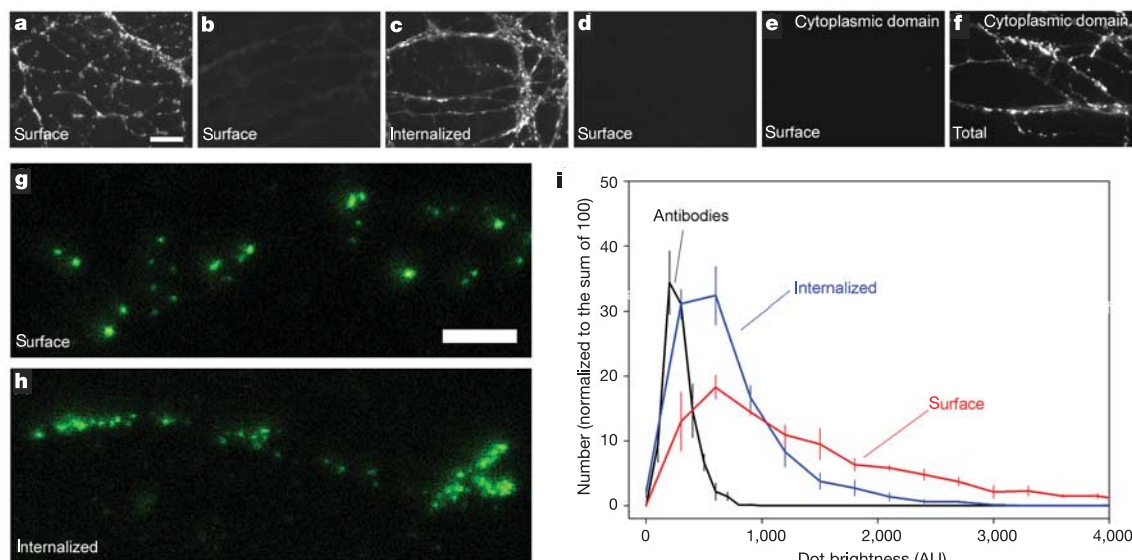


Figure 2 | Comparison between surface-exposed and internalized pools of synaptotagmin shows that the protein remains clustered in the presynaptic plasma membrane. **a–f**, Widefield images of hippocampal neurons labelled with antibodies specific for the luminal (**a–d**) or cytoplasmic (**e, f**) domain of synaptotagmin. **a**, Selective staining of the surface pool (low temperature and no calcium, to prevent endocytosis during labelling). **b**, Surface staining as in **a** but with secondary antibody added after blocking the surface epitopes of the bound monoclonal antibody using unlabelled anti-mouse antibodies (negative control). **c**, Selective staining of the internalized pool (labelled under conditions of active endocytosis followed by surface blocking as in **b**, but permeabilized after

fixation). **d**, Cultures stained with secondary antibodies only. **e, f**, Controls showing that an antibody specific for the cytoplasmic domain of synaptotagmin has no access to the epitope (**e**) unless the preparation is permeabilized after fixation (**f**). Scale bar, 10 μ m. **g**, STED image of surface-stained synapses (conditions as in **a**). Scale bar, 1 μ m. **h**, STED image of internalized vesicles (conditions as in **c**). **i**, Quantification of the brightness of single dots (in arbitrary units), compared with the dot brightness derived from single primary antibodies adsorbed to glass (see Methods). The graphs represent averages of 3–5 independent experiments ($\pm$ s.e.m.). Note that the bin size is 100 units for the single antibody graph and 300 units for the other images.

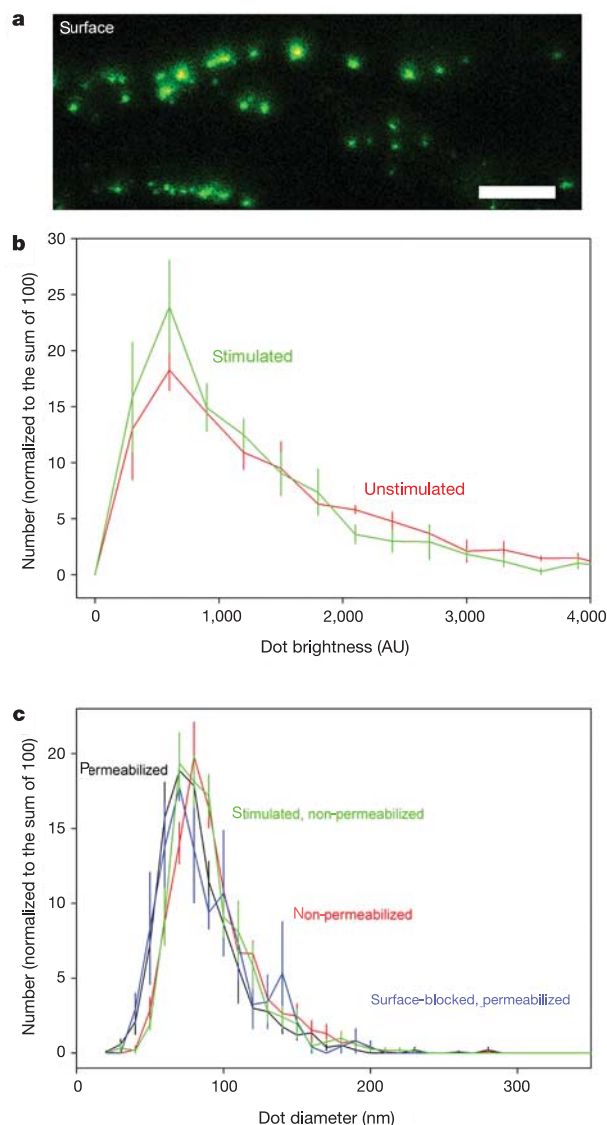


Figure 3 | Dot sizes of surface-exposed and internalized synaptotagmin pools do not show major differences, and neither dot brightness nor dot sizes change upon stimulation of exocytosis. **a**, Typical STED image of a heavily stimulated, surface-stained, non-permeabilized preparation. Compare with Fig. 2g. Scale bar, 1 μm . **b**, Comparison of dot brightness of the surface pool in stimulated and unstimulated preparations (mean $\pm$ s.e.m. from 4–5 independent experiments). **c**, Quantification of dot FWHM for different preparations (see Methods). We analysed the dots from 3–5 different experiments. To ensure that dot FWHM measurements were not affected by dots that were closely together, a stringent χ^2 cutoff of less than 0.01 (difference between the fit and the data) was placed on the lorentzian fit to the dots. See also Supplementary Fig. S2.

METHODS

Primary cultures. Primary cultures of hippocampal neurons were prepared from newborn rats (following procedures in refs 24, 25) and were used between days 15 and 25 *in vitro*. We used cultures plated onto astrocytic monolayers as well as cultures plated directly onto coverslips, with identical results. For antibody labelling, the neurons were incubated with a solution of anti-synaptotagmin monoclonal antibody (604.1, ref. 15) for 5 min at either 37 °C or on ice (for selective surface staining) in a Tyrode solution containing 124 mM NaCl, 5 mM KCl, 2 mM CaCl_2 , 1 mM MgCl_2 , 30 mM glucose and 25 mM HEPES (pH 7.4), washed briefly and fixed with 4% paraformaldehyde in PBS containing 1 mM EGTA. Where indicated, we stimulated the cultures at 37 °C in the presence of anti-synaptotagmin antibodies, in normal buffer containing calcium, before fixation. Surface-selective staining was then achieved by applying the secondary antibodies in the absence of permeabilization. After washing,

sheep anti-mouse Atto 532-labelled antibodies were added for 2 h at 22 °C; the cultures were then mounted in Mowiol (containing 6 g glycerol AR (4094, Merck), 2.4 g Mowiol 4-88 (Hoechst), 6 ml H_2O , 12 ml 0.2 M Tris pH 7.2 buffer) and imaged. Permeabilization was achieved using 0.1% Triton X-100 after fixation. The use of monoclonal (rather than polyclonal) antibodies ensures that any molecular patches observed cannot result from antibody-induced clustering. For investigation of the signal coming from single primary antibodies, coverslips treated with poly-L-lysine were incubated with anti-synaptotagmin antibodies for 30 min, fixed, washed, incubated with secondary antibodies and mounted. For the experiments in Fig. 2e, f, we used monoclonal antibody 41.1, which is directed against the cytoplasmic domain of synaptotagmin¹⁷.

Conventional imaging. Wide-field images were obtained with an epifluorescence microscope. The rhodamine dye Atto 532 was detected using a 480/20 excitation filter, a 530 DCLP beamsplitter and a 560/40 emission filter. For detection of Cy3 fluorescence (Fig. 2e, f) a 565/30 excitation filter, 595 DCLP beamsplitter and 645/745 emission filter were used (all filters were obtained from AHF Analysentechnik). Cy 3-labelled goat anti-mouse secondary antibodies were obtained from Jackson ImmunoResearch.

STED microscopy. STED microscopy was performed with a home-built setup. Dye excitation was accomplished with a laser diode (Picoquant) emitting 100-ps pulses. The diode was triggered by STED pulses delivered by an optical parametric oscillator (APE) that was synchronously pumped by a Ti:Sapphire laser (MaiTai, Spectra Physics) operating at 80 MHz. The STED pulses were stretched to 200 ps by dispersion in a glass fibre and converted into a focal doughnut by means of a spatial light modulator (Hamamatsu) delivering a helical phase ramp (0– 2π) with a central singularity²⁶. Both beams were coupled into an oil immersion lens (HCX PL APO, $\times 100$, Leica) using custom-designed dichroic mirrors. The average power of the excitation and STED beams at the sample was 2 μW and 14 mW, respectively. The fluorescence collected by the lens was imaged onto a counting avalanche photodiode with an opening diameter of 71% of that of the back-projected fluorescence spot. STED-imaging was obtained by piezo-scanning the sample at a pixel dwell time of ~ 0.3 ms. With a pixel size of $15 \times 15 \text{ nm}^2$, the recording of a $1 \times 1 \mu\text{m}^2$ area involved an effective exposure time of 1.4 s.

The spot size and hence the resolution Δr of a STED microscope follows a new law $\Delta r \approx \lambda / (2n \sin \alpha \sqrt{1 + I/I_{\text{sat}}})$, where λ and $n \sin \alpha$ denote the wavelength and numerical aperture of the lens, respectively^{11,13,23}. I is the maximal focal intensity applied for STED and I_{sat} is a characteristic value at which the fluorescence probability is reduced to $\sim 1/e$. Unlike in a confocal or epifluorescence light microscope, for $I/I_{\text{sat}} \rightarrow \infty$ it follows that $\Delta r \rightarrow 0$, meaning that the resolution is no longer limited by diffraction.

Fluorescence labelling. For the secondary antibody, the green-emitting dye Atto 532 (a gift from K. H. Drexhage) was coupled to an anti-mouse IgG (Jackson ImmunoResearch) via its succinimidyl ester.

Data analysis. Dot brightness and FWHM were evaluated with macros in MatLab (Mathworks). The user interactively defined a region enclosing the dot of fluorescence in the image. The background level and the FWHM of focal spots in the x - and y -direction were derived from lorentzian fits. The brightness was defined as the background-corrected sum over all pixels within the FWHM. Averages of the values for the dots in the x - and y -directions were retained for each dot. Graphs show histograms of FWHM and brightness as mean $\pm$ s.e.m.

Received 18 November 2005; accepted 20 January 2006.

1. Sudhof, T. C. The synaptic vesicle cycle. *Annu. Rev. Neurosci.* **27**, 509–547 (2004).
2. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated emission depletion microscopy. *Opt. Lett.* **19**, 780–782 (1994).
3. Ceccarelli, B., Hurlbut, W. P. & Mauro, A. Turnover of transmitter and synaptic vesicles at frog neuromuscular junction. *J. Cell Biol.* **57**, 499–524 (1973).
4. Heuser, J. E. & Reese, T. S. Evidence for recycling of synaptic vesicle membrane during transmitter release at frog neuromuscular junction. *J. Cell Biol.* **57**, 315–344 (1973).
5. Rizzoli, S. O. & Betz, W. J. Synaptic vesicle pools. *Nature Rev. Neurosci.* **6**, 57–69 (2005).
6. Royle, S. J. & Lagnado, L. Endocytosis at the synaptic terminal. *J. Physiol. (Lond.)* **553**, 345–355 (2003).
7. Gandhi, S. P. & Stevens, C. F. Three modes of synaptic vesicular recycling revealed by single-vesicle imaging. *Nature* **423**, 607–613 (2003).
8. Aravanis, A. M., Pyle, J. L. & Tsien, R. W. Single synaptic vesicles fusing transiently and successively without loss of identity. *Nature* **423**, 643–647 (2003).
9. Zenisek, D., Steyer, J. A., Feldman, M. E. & Almers, W. A membrane marker

- leaves synaptic vesicles in milliseconds after exocytosis in retinal bipolar cells. *Neuron* **35**, 1085–1097 (2002).
10. Klar, T. A., Jakobs, S., Dyba, M., Egner, A. & Hell, S. W. Fluorescence microscopy with diffraction resolution limit broken by stimulated emission. *Proc. Natl Acad. Sci. USA* **97**, 8206–8210 (2000).
 11. Hell, S. W. Toward fluorescence nanoscopy. *Nature Biotechnol.* **21**, 1347–1355 (2003).
 12. Hell, S. W. in *Topics in Fluorescence Spectroscopy* (ed. Lakowicz, J. R.) 361–422 (Plenum Press, New York, 1997).
 13. Hell, S. W. Strategy for far-field optical imaging and writing without diffraction limit. *Phys. Lett. A* **326**, 140–145 (2004).
 14. Wilson, T. & Sheppard, C. J. R. *Theory and Practice of Scanning Optical Microscopy* (Academic, New York, 1984).
 15. Chapman, E. R. & Jahn, R. Calcium-dependent interaction of the cytoplasmic region of synaptotagmin with membranes. Autonomous function of a single C2-homologous domain. *J. Biol. Chem.* **269**, 5735–5741 (1994).
 16. Kraszewski, K. *et al.* Synaptic vesicle dynamics in living cultured hippocampal neurons visualized with Cy3-conjugated antibodies directed against the luminal domain of synaptotagmin. *J. Neurosci.* **15**, 4328–4342 (1995).
 17. Brose, N., Petrenko, A. G., Sudhof, T. C. & Jahn, R. Synaptotagmin: a calcium sensor on the synaptic vesicle surface. *Science* **256**, 1021–1025 (1992).
 18. Pyle, J. L., Kavalali, E. T., Piedras-Renteria, E. S. & Tsien, R. W. Rapid reuse of readily releasable pool vesicles at hippocampal synapses. *Neuron* **28**, 221–231 (2000).
 19. Harris, K. M. & Sultan, P. Variation in the number, location and size of synaptic vesicles provides an anatomical basis for the nonuniform probability of release at hippocampal CA1 synapses. *Neuropharmacology* **34**, 1387–1395 (1995).
 20. Sankaranarayanan, S. & Ryan, T. A. Real-time measurements of vesicle-SNARE recycling in synapses of the central nervous system. *Nature Cell Biol.* **2**, 197–204 (2000).
 21. Poskanzer, K. E., Marek, K. W., Sweeney, S. T. & Davis, G. W. Synaptotagmin I is necessary for compensatory synaptic vesicle endocytosis *in vivo*. *Nature* **426**, 559–563 (2003).
 22. Bennett, M., Calakos, N., Kreiner, T. & Scheller, R. Synaptic vesicle membrane proteins interact to form a multimeric complex. *J. Cell Biol.* **116**, 761–775 (1992).
 23. Westphal, V. & Hell, S. W. Nanoscale resolution in the focal plane of an optical microscope. *Phys. Rev. Lett.* **94**, 143903 (2005).
 24. Rosenmund, C. & Stevens, C. F. The rate of aldehyde fixation of the exocytotic machinery in cultured hippocampal synapses. *J. Neurosci. Methods* **76**, 1–5 (1997).
 25. Klingauf, J., Kavalali, E. T. & Tsien, R. W. Kinetics and regulation of fast endocytosis at hippocampal synapses. *Nature* **394**, 581–585 (1998).
 26. Török, P. & Munro, P. R. T. The use of Gauss-Laguerre vector beams in STED microscopy. *Opt. Expr.* **12**, 3605–3617 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors thank E. Neher for helpful comments. S.O.R. acknowledges fellowships from the European Molecular Biology Organization and from the Human Frontier Science Program. This work was partly supported by a grant from the Leibniz Program of the Deutsche Forschungsgemeinschaft awarded to R.J., a grant from the German Ministry of Research and Education to S.W.H., and by the DFG-Centre for Molecular Physiology of the Brain. We thank I. Herefort, M. Wienisch and J. Klingauf for assistance with cell culturing, and A. Schönle for help with his software ImSpector.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.J. (rjahn@gwdg.de).

LETTERS

Production of the antimalarial drug precursor artemisinic acid in engineered yeast

Dae-Kyun Ro^{1*}, Eric M. Paradise^{2*}, Mario Ouellet¹, Karl J. Fisher⁶, Karyn L. Newman¹, John M. Ndungu³, Kimberly A. Ho¹, Rachel A. Eachus¹, Timothy S. Ham⁴, James Kirby², Michelle C. Y. Chang¹, Sydnor T. Withers², Yoichiro Shiba², Richmond Sarpong³ & Jay D. Keasling^{1,2,4,5}

Malaria is a global health problem that threatens 300–500 million people and kills more than one million people annually¹. Disease control is hampered by the occurrence of multi-drug-resistant strains of the malaria parasite *Plasmodium falciparum*^{2,3}. Synthetic antimalarial drugs and malarial vaccines are currently being developed, but their efficacy against malaria awaits rigorous clinical testing^{4,5}. Artemisinin, a sesquiterpene lactone endoperoxide extracted from *Artemisia annua* L (family Asteraceae; commonly known as sweet wormwood), is highly effective against multi-drug-resistant *Plasmodium* spp., but is in short supply and unaffordable to most malaria sufferers⁶. Although total synthesis of artemisinin is difficult and costly⁷, the semi-synthesis of artemisinin or any derivative from microbially sourced artemisinic acid, its immediate precursor, could be a cost-effective, environmentally friendly, high-quality and reliable source of artemisinin^{8,9}. Here we report the engineering of *Saccharomyces cerevisiae* to produce high titres (up to 100 mg l⁻¹) of artemisinic acid using an engineered mevalonate pathway, amorphaadiene synthase, and a novel cytochrome P450 monooxygenase (CYP71AV1) from *A. annua* that performs a three-step oxidation of amorpha-4,11-diene to artemisinic acid. The synthesized artemisinic acid is transported out and retained on the outside of the engineered yeast, meaning that a simple and inexpensive purification process can be used to obtain the desired product. Although the engineered yeast is already capable of producing artemisinic acid at a significantly higher specific productivity than *A. annua*, yield optimization and industrial scale-up will be required to raise artemisinic acid production to a level high enough to reduce artemisinin combination therapies to significantly below their current prices.

We engineered artemisinic-acid-producing yeast in three steps, by (1) engineering the farnesyl pyrophosphate (FPP) biosynthetic pathway to increase FPP production and decrease its use for sterols, (2) introducing the amorphaadiene synthase gene (*ADS*) from *A. annua* into the high FPP producer to convert FPP to amorphaadiene, and (3) cloning a novel cytochrome P450 that performs a three-step oxidation of amorphaadiene to artemisinic acid from *A. annua* and expressing it in the amorphaadiene producer (Fig. 1). The first committed reaction in artemisinin biosynthesis is catalysed by ADS¹⁰, which has been characterized and used for *de novo* production of amorphaadiene in *Escherichia coli*¹¹. To test for improvements in FPP production, we expressed *ADS* under the control of the *GALI* promoter on the pRS425 plasmid (see Supplementary Information for details). Yeast engineered with *ADS* alone produced a low quantity of amorphaadiene (Fig. 2, strain EPY201, 4.4 mg l⁻¹ amorphaadiene).

To increase FPP production in *S. cerevisiae*, the expression of several genes responsible for FPP synthesis was upregulated, and one gene responsible for FPP conversion to sterols was downregulated. All of these modifications to the host strain were made by chromosomal integration to ensure the genetic stability of the host strain. Overexpression of a truncated, soluble form of 3-hydroxy-3-methylglutaryl-coenzyme A reductase (*tHMGR*)¹² improved amorphaadiene production approximately fivefold (Fig. 2, strain EYP208). Down-regulation of *ERG9*, which encodes squalene synthase (the first step after FPP in the sterol biosynthetic pathway), using a methionine-repressible promoter (*P_{MET3}*)¹³ increased amorphaadiene production an additional twofold (Fig. 2, strain EPY225). Although *upc2-1*, a semi-dominant mutant allele that enhances the activity of *UPC2* (a global transcription factor regulating the biosynthesis of sterols in *S. cerevisiae*)¹⁴, had only a modest effect on amorphaadiene production when overexpressed in the EPY208 background (Fig. 2, strain EPY210), the combination of downregulating *ERG9* and overexpressing *upc2-1* increased amorphaadiene production to 105 mg l⁻¹ (Fig. 2, strain EPY213). Integration of an additional copy of *tHMGR* into the chromosome further increased amorphaadiene production by 50% to 149 mg l⁻¹ (Fig. 2, strain EPY219). Although overexpression of the gene encoding FPP synthase (*ERG20*) had little effect on total amorphaadiene production (Fig. 2, strain EPY224), the specific production increased by about 10% owing to a decrease in cell density. Combining all of these modifications resulted in a strain (EPY224) able to produce 153 mg l⁻¹ amorphaadiene, a sesquiterpene production level nearly 500-fold higher than previously reported¹⁵.

To create a strain that produced artemisinic acid from amorphaadiene, we isolated genes encoding enzymes responsible for oxidizing amorphaadiene to artemisinic acid in *A. annua*. Artemisinin is a sesquiterpene lactone derivative, which is the most widespread and characteristic class of secondary metabolites found in Asteraceae (also known as Compositae)¹⁶. We hypothesized that plants belonging to the Asteraceae family would share common ancestor enzymes for the early steps in the biosynthesis of sesquiterpene lactones, and therefore undertook a comparative genomic analysis of plants in the Asteraceae family. Previous cell-free assays have indicated that a cytochrome P450 monooxygenase (P450) catalyses the first regio-specific hydroxylation of amorphaadiene (Fig. 1) in *A. annua*¹⁷. We thus retrieved P450-expressed-sequence tags (ESTs) from the Asteraceae EST-database generated from two Asteraceae crops, sunflower and lettuce (<http://www.cgpdb.ucdavis.edu>). Use of degenerate primers highly specific to the Asteraceae CYP71 and CYP82 subfamilies (the most abundant P450 subfamilies in Asteraceae) enabled the isolation

¹California Institute of Quantitative Biomedical Research, ²Department of Chemical Engineering, ³Department of Chemistry, ⁴Department of Bioengineering, and ⁵Berkeley Center for Synthetic Biology, Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA. ⁶Amyris Biotechnologies Inc., Emeryville, California 94608, USA.

*These authors contributed equally to this work.

of several unique P450 fragments from an *A. annua* trichome-enriched complementary DNA pool. Using BLAST¹⁸ analyses of these P450 gene fragments against sunflower and lettuce ESTs, we identified a single *A. annua* P450 gene fragment that had surprisingly high sequence identity (85–88% at the amino-acid level) to ESTs of unknown function from both sunflower and lettuce. Sequence identity of this *A. annua* P450 fragment to other P450 fragments outside the Asteraceae family was much lower (~50% at the amino-acid level), indicating that this P450 is highly conserved in three distantly related genera in the Asteraceae family, but not in plants outside the Asteraceae family. This P450 gene was therefore a good candidate for a conserved Asteraceae sesquiterpene lactone biosynthetic enzyme.

The corresponding full-length P450 cDNA (*CYP71AV1*), encoding an open reading frame of 495 amino acids, was recovered from *A. annua*. Phylogenetic analysis showed that *CYP71AV1* shares a close lineage with other P450s that catalyse the hydroxylation of

monoterpenoids (*CYP71D13/18*; ref. 19), sesquiterpenoids (*CYP71D20*; ref. 20) or diterpenoids (*CYP71D16*; ref. 21), further suggesting the potential involvement of this P450 in terpenoid metabolism (Supplementary Information). For functional, heterologous expression of *CYP71AV1*, its native redox partner, NADPH: cytochrome P450 oxidoreductase (*CPR*), was also isolated from *A. annua*, and its biochemical function was confirmed *in vitro*. (Michaelis–Menten constants (K_m) for cytochrome *c* and NADPH were determined to be $4.3 \pm 0.7 \mu\text{M}$ and $23.0 \pm 4.4 \mu\text{M}$ (mean $\pm$ s.d., $n = 3$), respectively.)

Using *A. annua* *CPR* as a redox partner for *CYP71AV1*, we then investigated whether *CYP71AV1* could catalyse the conversion of amorpha-4,11-diene to more oxygenated products *in vivo*. The transgenic yeast strain EPY224 was transformed with a vector harbouring *CPR* and *CYP71AV1* under the control of galactose-inducible promoters. After galactose induction, the ether-extractable fraction of the yeast culture medium and cell pellet were analysed by gas chromatography mass spectrometry (GC–MS). A single chromatographic peak unique to EPY224 co-expressing *CYP71AV1* and *CPR* was detected in both the yeast culture medium and cell pellet, but was not present in control yeast (EPY224 expressing *CPR* only). However, more than 95% of this novel compound was associated with the cell pellet. In GC–MS analysis, the electron-impact mass spectrum and retention time of this compound were identical to those of artemisinic acid isolated from *A. annua* (Fig. 3). In a shake-flask culture, $32 \pm 13 \text{ mg l}^{-1}$ (mean $\pm$ s.d., $n = 7$) artemisinic acid was produced from EPY224 expressing *CYP71AV1* and *CPR*. Notably, the pathway intermediates, artemisinic alcohol and artemisinic aldehyde, were present at negligible levels in the culture medium and cell pellets of EPY224 engineered with *CYP71AV1* and *CPR*. (Artemisinic alcohol was present at less than 5% of the artemisinic acid in the cell pellet, and no artemisinic aldehyde was detected.)

Almost all (>96%) of the synthesized artemisinic acid was removed from the cell pellet by washing with alkaline buffer (pH 9 Tris-HCl buffer supplemented with 1.2 M sorbitol), with less than 2% remaining in the washed cell pellet or culture medium. Thus, it seems that artemisinic acid is efficiently transported out of yeast cells but remains bound to the cell surface when it is protonated under acidic culture conditions. We used this feature to develop a one-step purification method: a single silica gel column chromatographic separation of ether-extracted artemisinic acid from the wash buffer routinely yielded >95% pure artemisinic acid. In a one-litre aerated bioreactor, 115 mg of artemisinic acid was produced, of which 76 mg was purified using this method. The ¹H and ¹³C nuclear magnetic resonance spectra of this yeast-derived artemisinic acid were identical to those of artemisinic acid isolated from the leaves of *A. annua*, and are consistent with previously reported values^{22,23}. We can therefore confirm that structurally authentic artemisinic acid is synthesized by

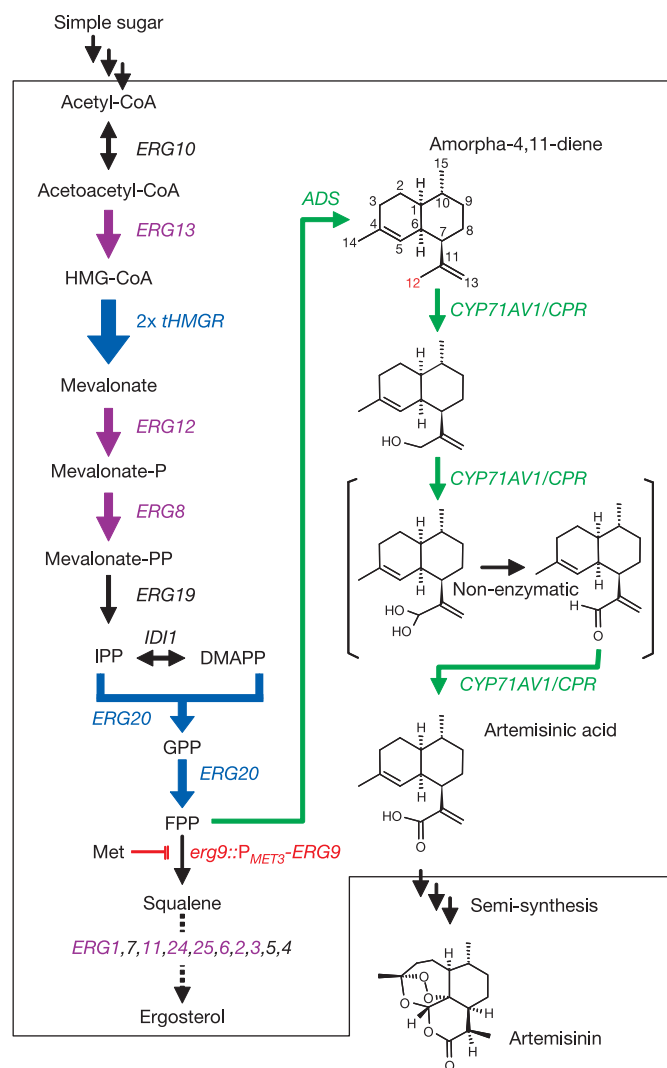


Figure 1 | Schematic representation of the engineered artemisinic acid biosynthetic pathway in *S. cerevisiae* strain EPY224 expressing *CYP71AV1* and *CPR*. Genes from the mevalonate pathway in *S. cerevisiae* that are directly upregulated are shown in blue; those that are indirectly upregulated by *upc2-1* expression are in purple; and the red line denotes repression of *ERG9* in strain EPY224. The pathway intermediates IPP, DMAPP and GPP are defined as isopentenyl pyrophosphate, dimethyl allyl pyrophosphate and geranyl pyrophosphate, respectively. Green arrows indicate the biochemical pathway leading from farnesyl pyrophosphate (FPP) to artemisinic acid, which was introduced into *S. cerevisiae* from *A. annua*. The three oxidation steps converting amorpha-4,11-diene to artemisinic acid by *CYP71AV1* and *CPR* are shown.

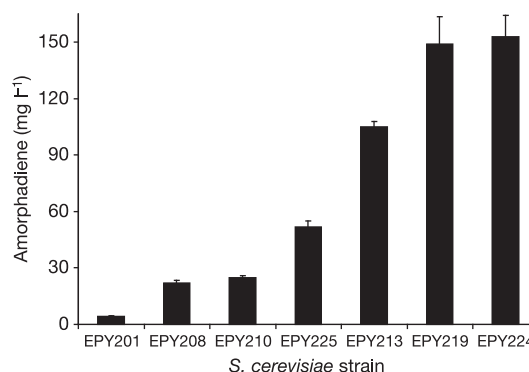


Figure 2 | Production of amorpha-4,11-diene by *S. cerevisiae* strains. The various *S. cerevisiae* strains are described in the text. Cultures were sampled after 144 h of growth, and amorpha-4,11-diene levels were quantified. Data, shown as total production, are mean $\pm$ s.d. ($n = 3$).

transgenic yeast *de novo*. The transgenic yeast produced artemisinic acid at a biomass fraction comparable to that produced by *A. annua* (4.5% dry weight in yeast compared to 1.9% artemisinic acid and 0.16% artemisinin in *A. annua*) but over a much shorter time (4–5 days for yeast versus several months for *A. annua*). As such, the specific productivity of the engineered yeast strain is nearly two orders of magnitude greater than *A. annua*.

Three-step oxidations by P450 enzymes have been previously reported in plant hormone gibberellin biosynthetic pathways^{24,25}. We conducted *in vitro* enzyme assays to identify whether CYP71AV1 catalyses all three oxidation reactions from amorphaadiene to artemisinic acid. Microsomes were isolated from *S. cerevisiae* strain YPH499 expressing either CPR alone or CPR and CYP71AV1, and incubated with pathway intermediates (amorphaadiene, artemisinic alcohol or artemisinic aldehyde) (Fig. 4). Microsomes from the CPR control did not catalyse the conversion of any pathway intermediate to more oxidized products, whereas efficient conversion of amorphaadiene, artemisinic alcohol and artemisinic aldehyde to the final product artemisinic acid was detected in microsomes containing

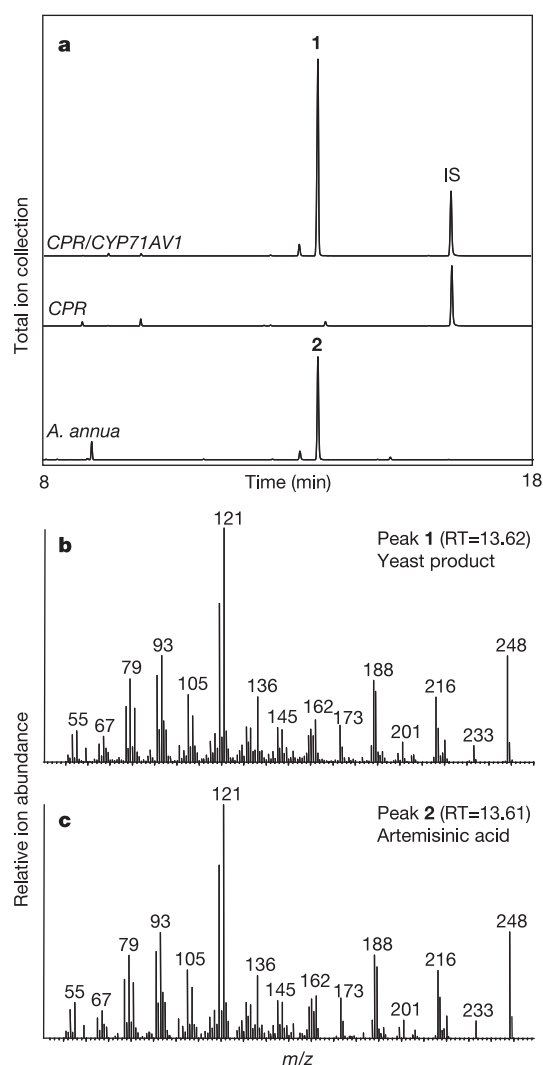


Figure 3 | GC-MS analysis of artemisinic acid produced from *A. annua* and transgenic yeast. **a**, Cell pellets from *S. cerevisiae* strain EPY224 expressing CPR or CPR and CYP71AV1 were washed using an alkaline buffer followed by acidification and ether extraction. Artemisinic acid was extracted from *A. annua* leaves using hexane. Methyl esters of both samples were prepared with trimethylsilyl-diazomethane before GC-MS analysis. The internal standard (IS) is the methyl ester of 4-octylbenzoic acid. **b**, **c**, Mass spectra and retention times of artemisinic acid from yeast (**b**) and *A. annua* (**c**). RT, retention time (in min).

CYP71AV1 and CPR. These *in vitro* assays demonstrate unambiguously that recombinant CYP71AV1 is able to catalyse three oxidation reactions at the C12 position of amorphaadiene. Previous *in vitro* enzyme assays using *A. annua* protein extract have suggested that soluble alcohol and aldehyde dehydrogenases and a C11,13 double-bond reductase (which acts on the aldehyde) are involved in artemisinin biosynthesis¹⁷. Although we cannot exclude a catalytic role for additional alcohol and aldehyde dehydrogenases in artemisinin synthesis in *A. annua*, the efficient *in vivo* conversion of amorpha-4,11-diene to artemisinic acid by recombinant CYP71AV1 indicates that the membrane-bound, multifunctional CYP71AV1 is a key contributor to artemisinin biosynthesis.

In summary, we have created a strain of *S. cerevisiae* capable of producing high levels of artemisinic acid by engineering the FPP biosynthetic pathway to increase FPP production and by expressing amorphaadiene synthase, a novel cytochrome P450 and its redox partner from *A. annua*. Given the existence of known, relatively high-yielding chemistry for the conversion of artemisinic acid to artemisinin or any other derivative that might be desired^{8,9}, microbially produced artemisinic acid is a viable source of this potent family of antimalarial drugs. Upon optimization of product titres, a conservative analysis suggests that artemisinin combination therapies could be offered significantly below their current prices (see Supplementary Information). In addition to cost savings, this bioprocess should not be subject to factors like weather or political climates that may affect plant cultivation. Furthermore, artemisinic acid from a microbial source can be extracted using an environmentally friendly process without worrying about contamination by other terpenes that are produced by plants, thereby increasing the ease with which it can be produced while reducing purification costs.

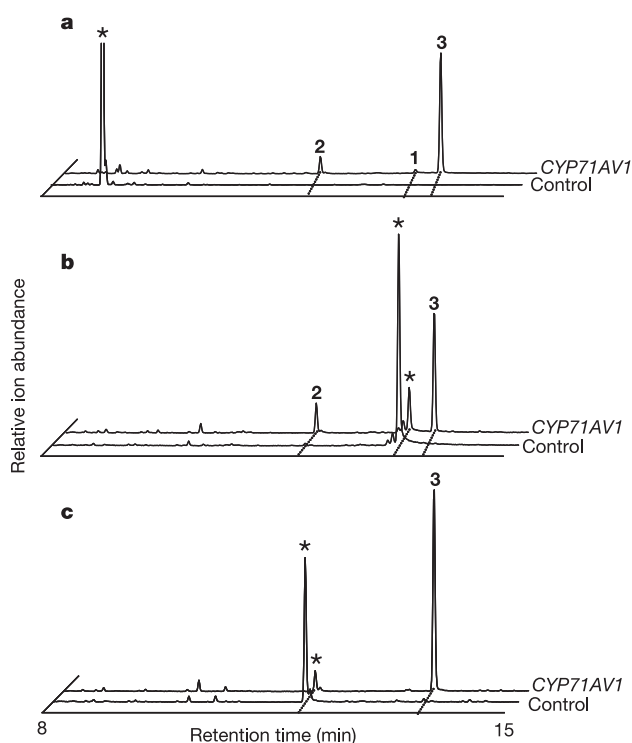


Figure 4 | *In vitro* enzyme assays for CYP71AV1 activities. Microsomes were isolated from *S. cerevisiae* strain YPH499 expressing CPR (control) or CPR and CYP71AV1 (CYP71AV1). **a–c**, For each enzyme assay, 10 μ M amorphaadiene (**a**), 25 μ M artemisinic alcohol (**b**) or 25 μ M artemisinic aldehyde (**c**) was used. Chromatographic peaks for the substrates used are indicated with asterisks. Ether-extractable fractions were derivatized and analysed by GC-MS in selective ion mode (m/z : 121, 189, 204, 218, 220 and 248). Enzymatic products are as indicated: 1, artemisinic alcohol (retention time 13.20 min); 2, artemisinic aldehyde (retention time 11.79 min); 3, artemisinic acid (retention time 13.58 min, detected as methyl ester).

METHODS

Detailed descriptions of the methods used in the generation and characterization of *S. cerevisiae* EPY strains, the cloning of *CYP71AV1* and *CPR*, and the semi-synthesis of artemisinic alcohol and artemisinic aldehyde are provided in Supplementary Information.

Chemicals and plant material. Authentic artemisinic acid was purchased from Apin Chemicals or extracted from *A. annua* leaves with hexane, as described in Supplementary Information. *A. annua* plants were started from seeds (Sandeman Seeds) and grown in a greenhouse at the University of California, Berkeley.

GC–MS analysis of amorphadiene. Amorphadiene production by the various strains was measured by GC–MS using a dodecane layer to trap volatile amorphadiene (see Supplementary Information for details). Amorphadiene (90% pure) was prepared by fermentation using an *E. coli* strain¹¹, and was used to construct a standard curve to determine amorphadiene production levels.

In vivo production, purification and chemical analysis of artemisinic acid. Pre-cultured EPY224 strains transformed with pESCURA::CPR or pESCURA::CPR/CYP71AV1 were inoculated at an absorbance of 0.05 at 600 nm (A_{600}) in 25 ml synthetic defined medium lacking histidine, leucine, methionine and uracil, and supplemented with 0.2% dextrose, 1.8% galactose and 1 mM methionine. After 120 h of culture at 30 °C, the cells were centrifuged and the cell pellet was washed using 50 mM Tris-HCl buffer (pH 9). The buffer was acidified to pH 2 using 2 M HCl, and extracted with ethyl acetate spiked with 4-octyl benzoic acid ($10 \mu\text{g ml}^{-1}$). The extracts were derivatized by 50 μl of 2 M TMS-diazomethane (Aldrich) with 10% methanol. For qualitative analysis by GC–MS, the product was purified by silica gel column chromatography eluted with ether and pentane (1:1).

Products were analysed using a gas chromatography mass spectrometer (70 eV, Agilent Technologies) equipped with a DB5 capillary column (0.25 mm internal diameter $\times$ 0.25 μm $\times$ 30 m, I&W Scientific). The gas chromatography oven programme used was 80 °C (held for 2 min), 20 °C min^{-1} ramp to 140 °C, product separation by a 5 °C min^{-1} increment up to 220 °C. For quantification by gas chromatography–flame ionization detection, samples were analysed without column purification using the same gas chromatography oven programme.

Fermentation and product analyses. A one-litre bioreactor (New-Brunswick Scientific) was used to culture the transgenic yeast strain for 93 h at 30 °C. Yeast cells were induced with 2% galactose at an A_{600} of 1.7, and the final cell density reached an A_{600} of 5.0. The dissolved oxygen level was maintained at 40% by altering agitation speed from 100 to 500 r.p.m. and sparging air at 0.5 l min^{-1} . Artemisinic acid was removed from the cell pellet by an alkaline wash as before, and purified through a silica column eluted with 78% hexane, 20% ethyl acetate and 2% acetic acid. The structure of the isolated artemisinic acid (>95% purity) was analysed by ¹H and ¹³C NMR using a 500 MHz NMR spectrometer (Bruker DRX-500) in the College of Chemistry NMR Facility at the University of California.

In vitro enzyme assays. A one-litre culture of *S. cerevisiae* YPH499 transformed with pESCURA::CPR or pESCURA::CPR/CYP71AV1 was induced with 2% galactose for 24 h. Microsomes were purified by polyethylene glycol precipitation followed by an additional ultracentrifugation step to remove cytosolic protein contamination, as previously described²⁶. Approximately 500 μg of total microsomal protein was used in a 1-ml reaction containing 100 mM potassium phosphate buffer pH 7.5, 10 or 25 μM substrate, 100 μM NADPH and an NADPH regeneration system (5 mM glucose-6-phosphate and two units of glucose-6-phosphate dehydrogenase). Reactions were incubated for 2 h at 24 °C with gentle agitation, acidified to pH 2, and extracted with ethyl ether. Products were separated using the same gas chromatography oven programme as above. Selective ion mode (SIM), including six ions characteristic to the products (121, 189, 204, 218, 220 and 248), was used for detection.

Received 22 December 2005; accepted 9 February 2006.

- World Health Organization. *World Malaria Report 2005*. (WHO, Geneva, 2005).
- Korenromp, E. L., Williams, B. G., Gouws, E., Dye, C. & Snow, R. W. Measurement of trends in childhood malaria mortality in Africa: An assessment of progress toward targets based on verbal autopsy. *Lancet Infect. Dis.* **3**, 349–358 (2003).
- Marsh, K. Malaria disaster in Africa. *Lancet* **352**, 924 (1998).
- Hoffman, S. Save the children. *Nature* **430**, 940–941 (2004).
- Vennerstrom, J. L. et al. Identification of an antimalarial synthetic trioxolane drug development candidate. *Nature* **430**, 900–904 (2004).
- Enserink, M. Infectious diseases: Source of new hope against malaria is in short supply. *Science* **307**, 33 (2005).
- Schmid, G. & Hofheinz, W. Total synthesis of Qinghaosu. *J. Am. Chem. Soc.* **105**, 624–625 (1983).
- Acton, N. & Roth, R. J. On the conversion of dihydroartemisinic acid into artemisinin. *J. Org. Chem.* **57**, 3610–3614 (1992).
- Haynes, R. K. & Vonwiller, S. C. Cyclic peroxyacetal lactone, lactol and

ether compounds. US patent 5,310,946 (1994).

- Mercke, P., Bengtsson, M., Bouwmeester, H. J., Posthumus, M. A. & Brodelius, P. E. Molecular cloning, expression, and characterization of amorphadiene synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. *Arch. Biochem. Biophys.* **381**, 173–180 (2000).
- Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D. & Keasling, J. D. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nature Biotechnol.* **21**, 796–802 (2003).
- Donald, K., Hampton, R. & Fritz, I. Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **63**, 3341–3344 (1997).
- Gardner, R. G. & Hampton, R. Y. A highly conserved signal controls degradation of 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase in eukaryotes. *J. Biol. Chem.* **274**, 31671–31678 (1999).
- Davies, B. S. J., Wang, H. S. & Rine, J. Dual activators of the sterol biosynthetic pathway of *Saccharomyces cerevisiae*: Similar activation/regulatory domains but different response mechanisms. *Mol. Cell. Biol.* **25**, 7375–7385 (2005).
- Jackson, B. E., Hart-Wells, E. A. & Matsuda, S. P. T. Metabolic engineering to produce sesquiterpenes in yeast. *Org. Lett.* **5**, 1629–1632 (2003).
- Seaman, F. C. Sesquiterpene lactones as taxonomic characters in the Asteraceae. *Bot. Rev.* **48**, 121–592 (1982).
- Bertea, C. M. et al. Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. *Planta Med.* **71**, 40–47 (2005).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- Lupien, S., Karp, F., Wildung, M. & Croteau, R. Regiospecific cytochrome P450 limonene hydroxylases from mint (*Mentha*) species: cDNA isolation, characterization, and functional expression of (–)-4S-limonene-3-hydroxylase and (–)-4S-limonene-6-hydroxylase. *Arch. Biochem. Biophys.* **368**, 181–192 (1999).
- Ralston, L. et al. Cloning, heterologous expression, and functional characterization of 5-*epi*-aristolochene-1,3-dihydroxylase from tobacco (*Nicotiana tabacum*). *Arch. Biochem. Biophys.* **393**, 222–235 (2001).
- Wang, E. et al. Suppression of a P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid resistance. *Nature Biotechnol.* **19**, 371–374 (2001).
- Elmarakby, S. A., El-Ferali, F. S., Elshohly, H. N., Croom, E. M. & Hufford, C. D. Microbiological transformations of artemisinic acid. *Phytochemistry* **27**, 3089–3091 (1988).
- Kim, S., Han, J. & Lim, Y. Revised assignment of ¹H-NMR signals of artemisinic acid. *Planta Med.* **62**, 480–481 (1996).
- Helliwell, C. A., Poole, A., Peacock, W. J. & Dennis, E. S. *Arabidopsis* ent-kaurene oxidase catalyzes three steps of gibberellin biosynthesis. *Plant Physiol.* **119**, 507–510 (1999).
- Helliwell, C. A., Chandler, P. M., Poole, A., Dennis, E. S. & Peacock, W. J. The CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc. Natl Acad. Sci. USA* **98**, 2065–2070 (2001).
- Pompon, D., Louerat, B., Bronine, A. & Urban, P. Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* **272**, 51–64 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Nelson for assigning an official CYP number, and D. Ockey and D. McPhee for artemisinic acid isolation from the *A. annua* plant. We are also grateful to N. A. Da Silva for p8-UB, R. Y. Hampton for pRH127-3 and pRH973, and J. Rine for pBD33 and pBD36. We thank R. Michelsmore and other members in the Compositae Genomics Project for the support of this project. This research was conducted under the sponsorship of the Institute for OneWorld Health, through the generous support of The Bill and Melinda Gates Foundation, and through funding from the Akibene Foundation, the United States Department of Agriculture, a University of California Discovery Grant, the Diversa Corporation and the National Science Foundation.

Author Contributions D.-K.R., E.M.P. and J.D.K. designed the project and experiments. D.-K.R., E.M.P., Y.S., M.C.Y.C., S.T.W. and J.K. performed experiments. K.J.F. conducted NMR analysis of artemisinic acid. J.M.N. and R.S. semi-synthesized artemisinic alcohol and artemisinic aldehyde. T.S.H. performed bioinformatics analysis of the Compositae EST-database. M.O., R.A.E. and K.A.H. provided technical assistance. D.-K.R., E.M.P., K.L.N. and J.D.K. wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information *Artemisia annua* CYP71AV1 and CPR gene sequence information has been deposited in GenBank under accession numbers DQ268763 and DQ318192, respectively. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.D.K. (keasling@berkeley.edu).

LETTERS

Reactive oxygen species have a causal role in multiple forms of insulin resistance

Nicholas Houstis^{1,2,3}, Evan D. Rosen^{1,4} & Eric S. Lander^{1,2,3,5}

Insulin resistance is a cardinal feature of type 2 diabetes and is characteristic of a wide range of other clinical and experimental settings. Little is known about why insulin resistance occurs in so many contexts. Do the various insults that trigger insulin resistance act through a common mechanism? Or, as has been suggested¹, do they use distinct cellular pathways? Here we report a genomic analysis of two cellular models of insulin resistance, one induced by treatment with the cytokine tumour-necrosis factor- α and the other with the glucocorticoid dexamethasone. Gene expression analysis suggests that reactive oxygen species (ROS) levels are increased in both models, and we confirmed this through measures of cellular redox state. ROS have previously been proposed to be involved in insulin resistance, although evidence for a causal role has been scant. We tested this hypothesis in cell culture using six treatments designed to alter ROS levels, including two small molecules and four transgenes; all ameliorated insulin resistance to varying degrees. One of these treatments was tested in obese, insulin-resistant mice and was shown to improve insulin sensitivity and glucose homeostasis. Together, our findings suggest that increased ROS levels are an important trigger for insulin resistance in numerous settings.

Insulin resistance is a key feature of type 2 diabetes. It also occurs in such clinical settings as pregnancy, sepsis, cancer cachexia, obesity, starvation, acromegaly, burn trauma and metabolic syndrome, and in response to many experimental treatments *in vitro* and *in vivo*. We chose to study insulin resistance resulting from the treatment of 3T3-L1 adipocytes with either the inflammatory cytokine tumour-necrosis factor- α (mouse TNF) or the glucocorticoid dexamethasone. Both dexamethasone and TNF are well-validated experimental models of insulin resistance, and both have physiological relevance *in vivo*. Mice show impaired insulin sensitivity in response to TNF or dexamethasone treatment, and are protected from obesity-related insulin resistance by related physiological blockades (genetic ablation of TNF or the TNF receptor², or treatment with glucocorticoid antagonists³, respectively). Glucocorticoid treatment is also a frequent cause of insulin resistance in humans. Furthermore, elevated levels of TNF- α or glucocorticoids (or both) have been shown to be associated with insulin-resistant states such as obesity^{4,5}, cancer cachexia⁶, sepsis⁷, burn trauma⁸, pregnancy^{9,10}, metabolic syndrome¹¹ and starvation¹².

Despite these similarities between TNF and dexamethasone, their cellular response pathways are quite distinct: TNF signals through a cell-surface cytokine receptor, whereas dexamethasone signals through a nuclear hormone receptor. Moreover, TNF has pro-inflammatory properties, whereas dexamethasone is a prototypical anti-inflammatory agent. We reasoned that a powerful approach to understanding the cellular basis of insulin resistance would be to compare the effects of these two very different but

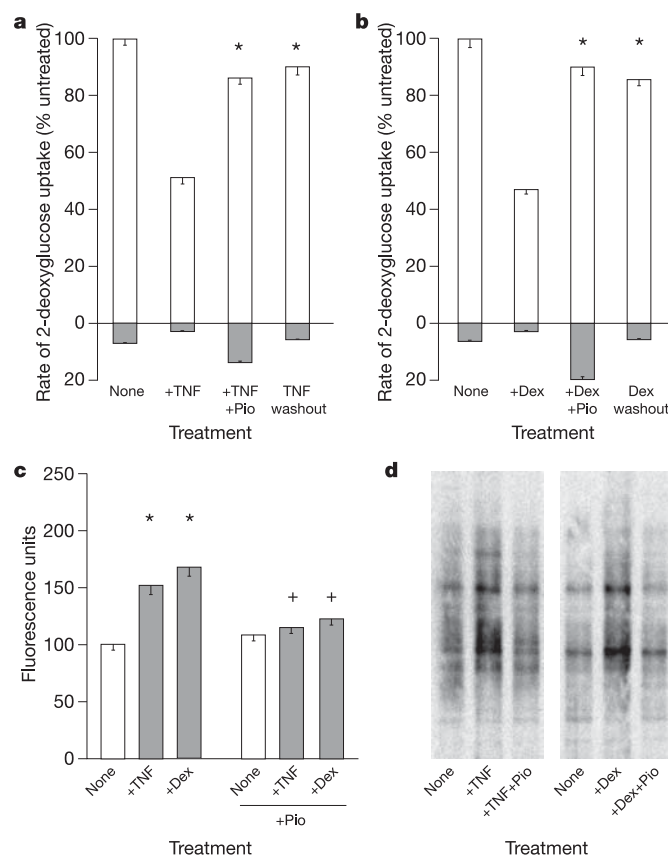


Figure 1 | Characterization of the insulin-resistant state. **a**, Rates of glucose transport in TNF-treated cells. Basal glucose transport (grey) and insulin-stimulated glucose transport (white) are shown. Cells were untreated, treated with TNF alone, TNF plus pioglitazone (Pio), or TNF followed by medium (TNF washout). Basal rate refers to the rate of glucose transport in the absence of insulin. Insulin-stimulated rate was calculated as the rate of transport in the presence of insulin minus the basal rate. All values are normalized to the insulin-stimulated rate from untreated cells. Asterisk indicates a significant difference ($P < 0.05$, t -test) compared to TNF treatment alone. **b**, Rates of glucose transport in dexamethasone (Dex)-treated cells. Data are analogous to **a**. **c**, Measurement of cellular redox status, showing rates of dichlorofluorescein (DCF) oxidation. Asterisk, $P < 0.05$ versus untreated; plus sign, $P < 0.05$ versus TNF or dexamethasone alone (t -tests). Results in **a–c** are mean $\pm$ s.e.m. **d**, Measurement of chronic oxidative stress. Immunoblots show total protein carbonylation. In the left panel, lanes 1–3 were loaded with an equal amount of protein from cells that were untreated, treated with TNF alone or treated with TNF plus pioglitazone. The right panel is analogous, with dexamethasone replacing TNF.

¹Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02141, USA. ²Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ³Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA. ⁴Division of Endocrinology, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts 02214, USA. ⁵Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA.

physiologically relevant treatments, as any pathways fundamental to insulin resistance might be expected to show responses in both settings.

Cultured 3T3-L1 adipocytes exposed to TNF or dexamethasone become insulin resistant within several days, as assessed by the ability of insulin to stimulate glucose uptake^{13,14}. To maximize the physiological relevance of this model, we calibrated the treatment regimen so that (1) insulin-dependent glucose uptake was decreased by ~50%, a degree similar to that seen in the clinical setting, (2) the defect in insulin action was reversible by washing out the agent, and (3) the defect could be rescued by pioglitazone (Fig. 1a, b), a member of the thiazolidinedione (TZD) class of insulin-sensitizing drugs.

We analysed genome-wide gene expression in TNF-treated, dexamethasone-treated and untreated adipocytes. Messenger RNA was prepared and hybridized to Affymetrix arrays containing probe sets from 22,690 mouse genes, with experiments performed in triplicate. Of the 350 upregulated genes showing the most robust expression change in response to TNF and dexamethasone treatments, only 34 of these were common to both treatments (Supplementary Table 1).

We analysed these results in two ways. First, we inspected the overlapping genes by eye, and noted that a substantial fraction (18%) was clearly related to the biology of reactive oxygen species (ROS). ROS are the radical forms of oxygen that arise as by-products of mitochondrial respiration and enzymatic oxidases. ROS are capable of acting as signalling molecules, but also cause damage to cellular

proteins, lipids and nucleic acids. The ROS-related genes with expression changes in response to both dexamethasone and TNF are shown in Supplementary Table 1.

Second, we applied the objective approach of gene set enrichment analysis (GSEA)^{15,16}. Given a collection of gene sets and the ranked list of gene expression changes resulting from a physiological treatment, GSEA tests whether the members of each set are randomly distributed along the list or clustered near the extremes, with the latter indicating that the gene set is regulated by the physiological treatment. From a curated collection of 475 gene sets¹⁶, GSEA identified ROS-related genes as the highest-scoring set for both dexamethasone treatment and TNF treatment (Supplementary Table 1b). Moreover, this was the only high-scoring set in common between the two treatments.

These gene expression results raise the possibility that ROS may be a key feature in both models of insulin resistance. We thus sought to test directly whether ROS levels are altered by TNF and dexamethasone treatment, by measuring oxidation of the redox-sensitive dye dichlorofluorescein (DCF). The resulting signal was higher by 50% or 65%, respectively, in insulin-resistant cells produced by TNF or dexamethasone treatment (Fig. 1c). Moreover, we found that increases in ROS levels precede the onset of detectable insulin resistance, becoming evident midway through treatment. In addition, we found that protein carbonyl levels, a marker of cumulative oxidative stress, were elevated by 50% and 110%, respectively,

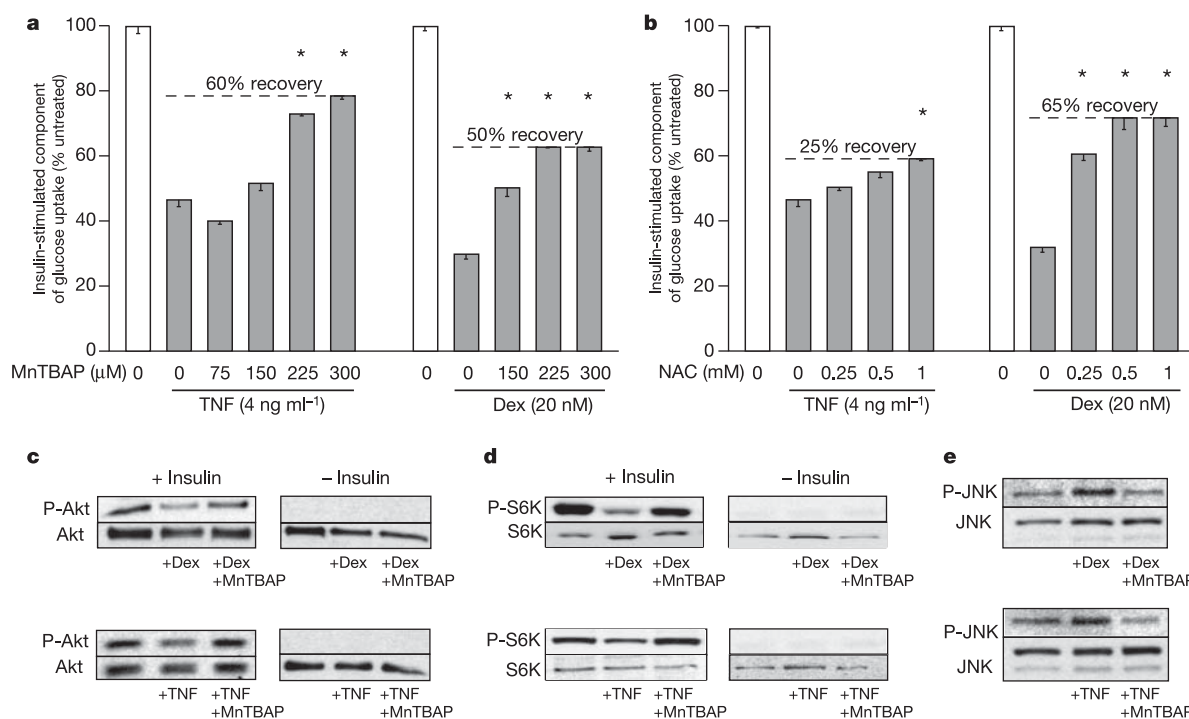


Figure 2 | Effect of antioxidants on insulin resistance. **a**, Partial restoration of TNF- or dexamethasone-induced insulin resistance by MnTBAP. Bars show insulin-dependent glucose transport. Results are mean $\pm$ s.e.m. Asterisks indicate $P < 0.05$ versus TNF or dexamethasone (Dex) treatment alone (t -test). **b**, Partial restoration of TNF- or dexamethasone-induced insulin resistance by NAC. Data are analogous to **a**. **c–e**, Effect of 250 μ M MnTBAP on measures of insulin signal transduction in cells with dexamethasone-induced (top) or TNF-induced (bottom) insulin resistance. **c**, Immunoblots show phospho-Akt (P-Akt) levels on acute stimulation of cells with insulin (left) or in the absence of insulin (right). Total Akt protein levels are shown in the panel immediately below the phosphorylated counterpart. Signals were quantified by densitometry and normalized to total protein levels. The level of insulin-stimulated phospho-Akt decreases by 40% and 60%, respectively, in TNF- and dexamethasone-treated cells.

MnTBAP treatment recovers phospho-Akt levels to within 5% (TNF) and 30% (dexamethasone) of that found in untreated cells. **d**, Stimulation of p70S6K phosphorylation (P-S6K) by insulin was assayed, with data arranged and analysed as in **c**. The level of insulin-stimulated phospho-S6K decreases by 35% and 65%, respectively, in TNF- and dexamethasone-treated cells. MnTBAP treatment recovers phospho-S6K levels to within 25% (both TNF and dexamethasone) of that found in untreated cells. **e**, Immunoblot showing JNK phosphorylation (P-JNK) levels on treatment with dexamethasone, dexamethasone and MnTBAP, TNF, or TNF and MnTBAP. Total JNK levels are shown in the panel immediately below phospho-JNK. The level of phospho-JNK increases by 25% and 80%, respectively, in TNF- and dexamethasone-treated cells. MnTBAP treatment restores phospho-JNK to untreated levels.

in TNF- and dexamethasone-treated cells. When the TZD pioglitazone was co-administered with either TNF or dexamethasone, the level of DCF oxidation and the extent of protein carbonylation were both near that observed in untreated cells (Fig. 1c, d).

We next sought to test whether ROS have a causal role in insulin resistance by assessing whether a variety of treatments chosen specifically as suppressors of ROS levels could also act as insulin sensitizers. We began by evaluating two small antioxidant molecules, *N*-acetylcysteine (NAC) and manganese (III) tetrakis (4-benzoic acid) porphyrin (MnTBAP). NAC stimulates the formation of the endogenous reducing agent glutathione, which cells use to scavenge H_2O_2 . MnTBAP has catalytic activities similar to the ROS-scavenging enzymes superoxide dismutase (SOD) and catalase; it protects mammalian cells from damage by H_2O_2 (ref. 17) and partially complements loss-of-function SOD mutations in bacteria and mice¹⁸. NAC and MnTBAP were applied to adipocytes concomitantly with TNF or dexamethasone. Both antioxidants showed dose-dependent suppression of insulin resistance induced by either treatment, preventing 25–65% of the defect in insulin-mediated glucose uptake (Fig. 2a, b); importantly, neither compound increased insulin action on its own (data not shown). Furthermore, MnTBAP largely prevented the increase in protein carbonylation (Supplementary Fig. 3).

We also examined the effect of MnTBAP on various parameters of insulin signalling. Treatment with TNF or dexamethasone decreases levels of insulin-stimulated serine phosphorylation on Akt and p70S6 kinases, whereas co-treatment with MnTBAP largely prevented this decrease (Fig. 2c, d). Increased ROS levels are known to stimulate threonine phosphorylation of JNK, a kinase previously linked to insulin resistance^{19,20}. TNF and dexamethasone treatment increased phosphorylation of JNK, whereas MnTBAP decreased it to nearly baseline levels (Fig. 2e).

We then explored the causal relationship between ROS and insulin

resistance by constructing four 3T3-L1 cell lines carrying transgenes encoding ROS-scavenging enzymes, including CuZnSOD, MnSOD, a form of catalase with its peroxisomal localization signal removed ('cytocalase'), and catalase targeted specifically to the mitochondrion ('mitocatalase') (see Methods). These transgenes were delivered by retroviral transduction, and expression was directed by an inducible promoter (Tet-on system) so as to minimize any effects of transgene expression on adipocyte differentiation. Transgene expression was induced before treatment with TNF or dexamethasone, resulting in a 3–5-fold increase in enzyme activity above endogenous levels (Supplementary Fig. 2). Mitocatalase and cytocalase were most potent, preventing up to 65% of the reduction in insulin-stimulated glucose uptake; CuZnSOD and MnSOD prevented 50% and 25%, respectively. Neither the transgene inducer (doxycycline) nor the inducible expression system itself had any effect on glucose transport (data not shown).

Together, our results using two small molecules and four transgenes indicate that diverse treatments chosen as suppressors of ROS can significantly reduce dexamethasone- and TNF-induced insulin resistance.

We next sought to extend these observations from cellular models to an *in vivo* model of insulin resistance, the leptin-deficient *ob/ob* mouse. These mice become extremely obese, developing significant insulin resistance and glucose intolerance by eight weeks of age. The mice are also known to exhibit markers of oxidative stress²¹. We tested whether chronic delivery of MnTBAP could improve glucose homeostasis. Male *ob/ob* mice received MnTBAP (at 2.5, 5 or 10 mg kg⁻¹ body weight), the TZD rosiglitazone (3 mg kg⁻¹ body weight) or vehicle alone. Treatments were administered daily, beginning at 8 weeks of age and continuing for a period of 12 weeks. The doses of MnTBAP are similar to or lower than doses shown to improve longevity in *Sod2* knockout mice¹⁸.

MnTBAP had no effect on the body weight of *ob/ob* mice over the

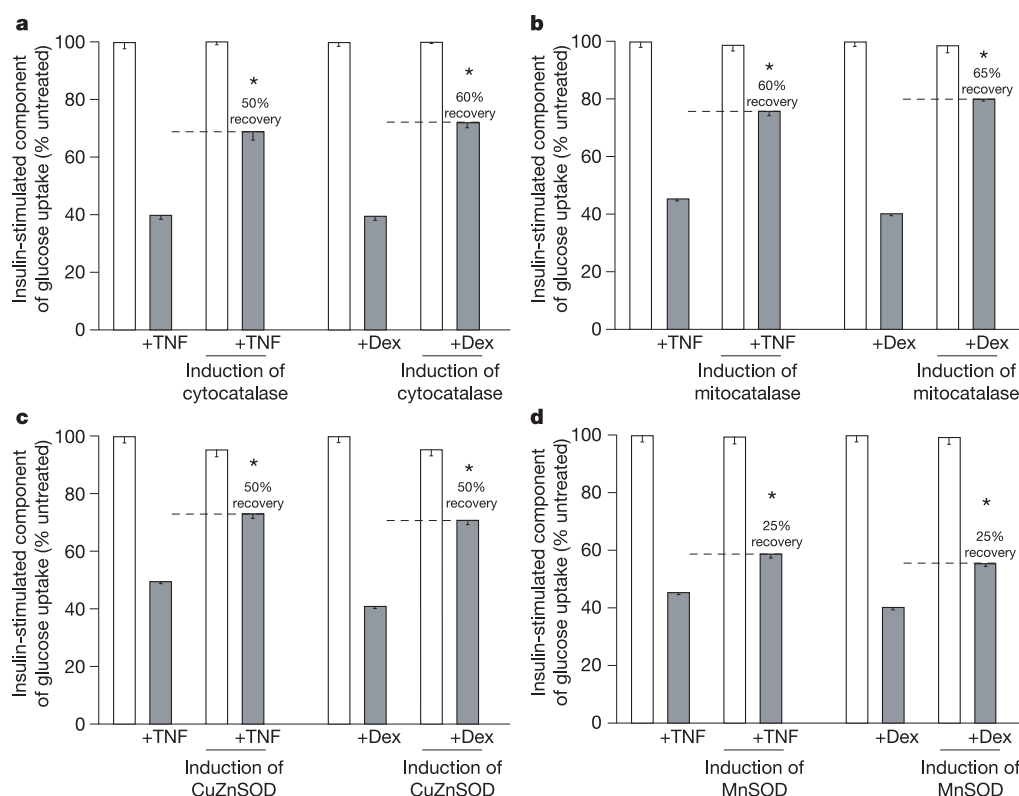


Figure 3 | Effect of transgenes on insulin resistance. a–d, Effect of cytocatalase (a), mitocatalase (b), CuZnSOD (c) or MnSOD (d) transgene expression on TNF- and dexamethasone-induced insulin resistance. Bars show insulin-dependent rates of glucose transport in untreated cells with or

without transgene induction (white), or in TNF- or dexamethasone-treated cells with or without transgene induction (grey). Results are mean $\pm$ s.e.m. Asterisks indicate $P < 0.05$ versus TNF or dexamethasone alone (*t*-test).

duration of the treatment period (Fig. 4a), but resulted in dose-dependent improvement in several measures of glucose homeostasis. For example, mice receiving the highest dose of MnTBAP showed nearly normal glucose levels in the fed state ($\sim 15\%$ above levels from control lean C57BL/6 mice in MnTBAP-treated versus $\sim 50\%$ in vehicle-treated mice) (Fig. 4b). MnTBAP improved glucose tolerance and insulin sensitivity (at 5 and 10 mg kg⁻¹), with a maximum effect comparable to that of animals treated with rosiglitazone (Fig. 4c, d). There was no significant effect on serum insulin levels in the fed or fasted state in MnTBAP-treated mice (data not shown).

In summary, these studies identify increased levels of ROS as a common feature of two models of insulin resistance, based on a genomic analysis that required no presupposition of which pathways might be involved. A causal role for ROS is shown by the observation that several pharmacological and genetic interventions designed to decrease ROS levels substantially prevent the development of insulin resistance. The treatment regimens tested in this study do not fully prevent the development of insulin resistance: this could either reflect that they do not restore ROS to normal levels, or could indicate that ROS levels act in parallel with other pathways. More detailed studies will be required.

ROS is one of many factors that have previously been suggested to

have a possible role in insulin resistance, on the basis of two types of indirect evidence: (1) extensive association of markers of oxidative stress with obesity and diabetes^{22,23}, and (2) experiments showing that direct treatment of 3T3-L1 adipocytes with high doses of hydrogen peroxide²⁴ or with agents that induce ROS accumulation²⁵ can induce insulin resistance. However, such findings do not imply that ROS have a causal role in any physiological models of insulin resistance. This would require evidence that insulin resistance can be prevented to some degree by blocking the increase in ROS levels. To our knowledge, the only previous results that bear on this point are experiments showing that treatment of rodents and humans with the compound α -lipoic acid can partially improve insulin sensitivity. However, this evidence has been difficult to interpret because the action of α -lipoic acid is unclear: it can act either as an antioxidant or a pro-oxidant, and can directly stimulate insulin-independent glucose uptake²⁶. We also note a recent study that has suggested that a chemical inhibitor of NADPH oxidase can improve glucose homeostasis in obese KKAY mice²², although its effects on insulin resistance *per se* were not studied.

Our findings lead to a number of predictions. For example, one might predict that clinical conditions associated with insulin resistance will also show evidence of increased ROS levels. In fact, markers of oxidative stress are elevated in sepsis, burn injury, starvation, obesity, acromegaly and type 2 diabetes. In addition, conditions that increase ROS levels would be predicted to cause insulin resistance. In a literature review, we identified four human diseases with primary defects that affect ROS balance: familial amyotrophic lateral sclerosis, Friedrich's ataxia, ataxia telangiectasia and acatalasia. In the first three disorders, metabolic studies have demonstrated significant insulin resistance, although no explicit connection with ROS has previously been drawn. In the fourth case, insulin resistance *per se* has not been characterized, but patients have high rates of diabetes. (Detailed references concerning the above clinical conditions are provided in Supplementary Information.)

These findings raise issues that require further investigation. First, what downstream pathways translate elevated ROS levels into insulin resistance? ROS have been shown to induce various signalling pathways involving FoxO, MAPK, JAK/STAT, p53, phospholipase C, PI(3)K and other proteins. The particular pathway activated will depend on the magnitude of oxidative stress, the specific type of ROS, the cell type, the duration of exposure, and other factors. One attractive possibility is that ROS-induced insulin resistance is mediated by JNK. JNK is known to be activated by oxidative stress²⁷, and inhibition of JNK activity (through genetic knockout¹⁹ or an inhibitory peptide²⁰) improves insulin sensitivity in mice. Our data show that JNK is activated in response to both TNF and dexamethasone, and that this effect is reversed by MnTBAP. Second, what is the source of the ROS in insulin resistance? It is unclear whether increased ROS levels in various settings arise through a common mechanism or through different pathways. One candidate for a common mechanism for both TNF and dexamethasone might be the sphingolipid ceramide²⁸, which is increased in TNF- and dexamethasone-treated cells as well as diabetic muscle, and is capable of inducing mitochondrial ROS formation²⁹ and insulin resistance in 3T3-L1 cells³⁰. Third, what adaptive role might be served by a cellular mechanism that decreases insulin sensitivity in response to elevated ROS levels? Because an imbalance between substrate availability and oxidative capacity can lead to increased levels of ROS, cells may interpret elevated ROS levels as a signal to limit the input to the electron transport chain by decreasing glucose uptake. Fourth, are there inducers of insulin resistance that act on pathways downstream of ROS and thus are not associated with increased levels of ROS? Finally, through which tissue(s) do ROS mediate their effects on whole-body insulin resistance? Although our *in vitro* data show that adipocytes are one possible site of oxidant action, we note that the use of MnTBAP *in vivo* might affect other organs as well, such as muscle or liver.

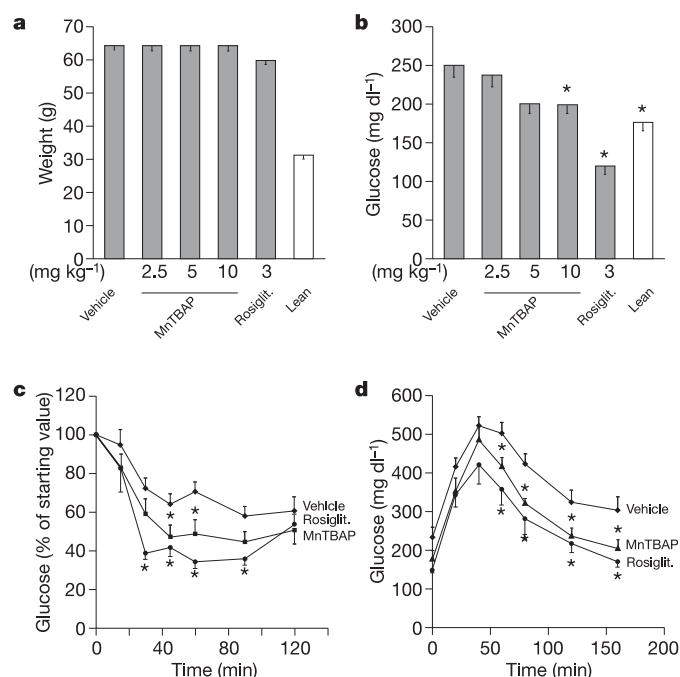


Figure 4 | Effects of chronic treatment with MnTBAP or rosiglitazone on obese mice. **a**, Effect on body weight. Bars show weights of animals after 12 weeks of treatment with daily subcutaneous injection of MnTBAP or rosiglitazone (Rosiglit.). Animals were treated with 2.5 mg kg⁻¹ ($n = 7$), 5 mg kg⁻¹ ($n = 8$) or 10 mg kg⁻¹ ($n = 8$) MnTBAP, 3 mg kg⁻¹ rosiglitazone ($n = 8$) or vehicle (phosphate-buffered saline) ($n = 8$). **b**, Effect on fed glucose levels. Bars indicate mean fed glucose levels averaged over 18 days during the last eight weeks of treatment. Glucose levels were determined at random times during the day and on random days during this period. **c**, Effect on insulin sensitivity. Insulin tolerance tests were performed on mice from each treatment group (vehicle, 5 mg kg⁻¹ MnTBAP or rosiglitazone). Lines indicate the time course of glucose levels after subcutaneous injection of human insulin (2 U kg⁻¹). **d**, Effect on glucose tolerance. Glucose tolerance tests were performed, with lines indicating the time course of glucose excursion following subcutaneous injection of glucose (1 g kg⁻¹) in vehicle-treated, 10 mg kg⁻¹ MnTBAP and rosiglitazone treatment groups. Results are mean $\pm$ s.e.m. Asterisks indicate $P < 0.05$ versus vehicle-treated animals (t -test). Glucose and insulin tolerance test data using different dose levels of MnTBAP are provided in Supplementary Information.

Finally, our results suggest that antioxidant therapy might be a useful strategy in type 2 diabetes and other insulin-resistant states. Although some agents, such as vitamin E, have been clinically tested with ambiguous results, these have tended to be weak antioxidants with limited ROS-reducing effects. Newer, more effective antioxidant agents now under development for use in atherosclerosis and neurodegenerative disorders may prove worthy of investigation in this regard.

METHODS

A more complete description of the methods used is provided in Supplementary Information.

Cell culture. Early passage 3T3-L1 pre-adipocytes were cultured in Dulbecco's Modified Eagle Medium supplemented with Glutamax, 10% bovine calf serum, 100 U ml⁻¹ penicillin and 0.1 mg ml⁻¹ streptomycin. Differentiation, glucose uptake assays and western blots were performed under standard conditions (see Supplementary Information).

Induction of insulin resistance. Treatment with dexamethasone (20 nM) or recombinant mouse TNF (4 ng ml⁻¹) was initiated with mature adipocytes anywhere from day 8 to day 14 of differentiation. Media was changed daily for TNF treatment, for a total incubation time of 4 days. Dexamethasone media was changed every other day for a total of 8 days. For experiments in which transgene expression was induced, doxycycline was added to the cells on day 4 of differentiation and TNF or dexamethasone treatment was always initiated on day 8.

ROS measurement. For DCF assays, cells were washed twice in KRP buffer, incubated in pre-warmed KRP buffer containing 25 mM glucose and 5 μ M CM-DCF, and placed at 37 °C. After 30 min, cells were washed once with KRP and fluorescence was immediately measured in a plate reader with an excitation/emission wavelength of 485/515 nm. DCF values were calculated after subtracting background fluorescence levels (measured under identical conditions but without DCF). For protein carbonylation assays, carbonyl levels were determined in 10 μ g adipocyte protein lysate using a commercial kit (OxyBlot, Serologicals), according to the manufacturer's instructions.

Animals. Male *ob/ob* mice were purchased from Jackson Laboratories and randomly assigned to treatment group, so as to ensure that each group had an equal average weight. Drugs were resuspended in PBS and each mouse received daily 300- μ l subcutaneous injections. Mice were weighed weekly.

Glucose and insulin tolerance tests. For the glucose tolerance test, mice were fasted for 12 h and then injected subcutaneously with glucose (1 g kg⁻¹ body weight). Blood samples were taken at regular time points (0–140 min), and blood glucose levels were determined with a portable glucose meter. For insulin tolerance tests, mice were fasted for 4 h and handled 30 min before performing the test. We then injected human regular insulin (2 U kg⁻¹ body weight) subcutaneously. Blood samples were taken at regular intervals (0–120 min) and blood glucose was measured as described above.

Received 23 November 2005; accepted 6 February 2006.

1. Kroder, G. *et al.* Tumor necrosis factor- α - and hyperglycemia-induced insulin resistance. Evidence for different mechanisms and different effects on insulin signaling. *J. Clin. Invest.* **97**, 1471–1477 (1996).
2. Uysal, K. T., Wiesbrock, S. M., Marino, M. W. & Hotamisligil, G. S. Protection from obesity-induced insulin resistance in mice lacking TNF- α function. *Nature* **389**, 610–614 (1997).
3. Kusunoki, M., Cooney, G. J., Hara, T. & Storlien, L. H. Amelioration of high-fat feeding-induced insulin resistance in skeletal muscle with the antigluco-corticoid RU486. *Diabetes* **44**, 718–720 (1995).
4. Bujalska, I. J., Kumar, S. & Stewart, P. M. Does central obesity reflect "Cushing's disease of the omentum"? *Lancet* **349**, 1210–1213 (1997).
5. Hotamisligil, G. S., Arner, P., Caro, J. F., Atkinson, R. L. & Spiegelman, B. M. Increased adipose tissue expression of tumor necrosis factor- α in human obesity and insulin resistance. *J. Clin. Invest.* **95**, 2409–2415 (1995).
6. McCall, J. L., Tuckey, J. A. & Parry, B. R. Serum tumour necrosis factor alpha and insulin resistance in gastrointestinal cancer. *Br. J. Surg.* **79**, 1361–1363 (1992).

7. Stoner, H. B. *et al.* The effect of sepsis on the oxidation of carbohydrate and fat. *Br. J. Surg.* **70**, 32–35 (1983).
8. Vaughan, G. M. *et al.* Cortisol and corticotrophin in burned patients. *J. Trauma* **22**, 263–273 (1982).
9. Hornnes, P. J. On the decrease of glucose tolerance in pregnancy. A review. *Diabete Metab.* **11**, 310–315 (1985).
10. Kirwan, J. P. *et al.* TNF- α is a predictor of insulin resistance in human pregnancy. *Diabetes* **51**, 2207–2213 (2002).
11. Wang, M. The role of glucocorticoid action in the pathophysiology of the Metabolic Syndrome. *Nutr. Metab. (Lond)* **2**, 3 (2005).
12. Balm, P. H. *Stress Physiology in Animals* (CRC Press, Boca Raton, 2000).
13. Hotamisligil, G. S., Murray, D. L., Choy, L. N. & Spiegelman, B. M. Tumor necrosis factor α inhibits signaling from the insulin receptor. *Proc. Natl Acad. Sci. USA* **91**, 4854–4858 (1994).
14. van Putten, J. P., Wieringa, T. & Krans, H. M. Corticosteroids as long-term regulators of the insulin effectiveness in mouse 3T3 adipocytes. *Diabetologia* **28**, 445–451 (1985).
15. Mootha, V. K. *et al.* PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genet.* **34**, 267–273 (2003).
16. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
17. Day, B. J., Fridovich, I. & Crapo, J. D. Manganic porphyrins possess catalase activity and protect endothelial cells against hydrogen peroxide-mediated injury. *Arch. Biochem. Biophys.* **347**, 256–262 (1997).
18. Melov, S. *et al.* A novel neurological phenotype in mice lacking mitochondrial manganese superoxide dismutase. *Nature Genet.* **18**, 159–163 (1998).
19. Hirosumi, J. *et al.* A central role for JNK in obesity and insulin resistance. *Nature* **420**, 333–336 (2002).
20. Kaneto, H. *et al.* Possible novel therapy for diabetes with cell-permeable JNK-inhibitory peptide. *Nature Med.* **10**, 1128–1132 (2004).
21. Yang, S. *et al.* Mitochondrial adaptations to obesity-related oxidant stress. *Arch. Biochem. Biophys.* **378**, 259–268 (2000).
22. Furukawa, S. *et al.* Increased oxidative stress in obesity and its impact on metabolic syndrome. *J. Clin. Invest.* **114**, 1752–1761 (2004).
23. Urakawa, H. *et al.* Oxidative stress is associated with adiposity and insulin resistance in men. *J. Clin. Endocrinol. Metab.* **88**, 4673–4676 (2003).
24. Rudich, A. *et al.* Prolonged oxidative stress impairs insulin-induced GLUT4 translocation in 3T3-L1 adipocytes. *Diabetes* **47**, 1562–1569 (1998).
25. Lin, Y. *et al.* The hyperglycemia-induced inflammatory response in adipocytes: the role of reactive oxygen species. *J. Biol. Chem.* **280**, 4617–4626 (2005).
26. Moini, H., Tirosh, O., Park, Y. C., Cho, K. J. & Packer, L. R- α -lipoic acid action on cell redox status, the insulin receptor, and glucose uptake in 3T3-L1 adipocytes. *Arch. Biochem. Biophys.* **397**, 384–391 (2002).
27. Kamata, H. *et al.* Reactive oxygen species promote TNF- α -induced death and sustained JNK activation by inhibiting MAP kinase phosphatases. *Cell* **120**, 649–661 (2005).
28. Summers, S. A. & Nelson, D. H. A role for sphingolipids in producing the common features of type 2 diabetes, metabolic syndrome X, and Cushing's syndrome. *Diabetes* **54**, 591–602 (2005).
29. Di Paola, M., Cocco, T. & Lorusso, M. Ceramide interaction with the respiratory chain of heart mitochondria. *Biochemistry* **39**, 6660–6668 (2000).
30. Summers, S. A., Garza, L. A., Zhou, H. & Birnbaum, M. J. Regulation of insulin-stimulated glucose transporter GLUT4 translocation and Akt kinase activity by ceramide. *Mol. Cell. Biol.* **18**, 5457–5464 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Zabolotny for assistance with signal transduction assays. We also thank D. Altshuler, O. Peroni and members of the Rosen and Lander laboratories for discussions. This work was supported in part by grants from the NIH (to E.D.R. and E.S.L.), the Hertz Foundation (N.H.) and the Broad Institute of MIT and Harvard.

Author Contributions E.D.R. and E.S.L. co-directed this work.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to E.D.R. (erosen@bidmc.harvard.edu) or E.S.L. (lander@broad.mit.edu).

I κ B kinase- α is critical for interferon- α production induced by Toll-like receptors 7 and 9

Katsuaki Hoshino¹, Takahiro Sugiyama¹, Mitsuru Matsumoto³, Takashi Tanaka¹, Masuyoshi Saito¹, Hiroaki Hemmi⁴, Osamu Ohara^{2,5}, Shizuo Akira⁴ & Tsuneyasu Kaisho¹

The Toll-like receptor (TLR) family has important roles in microbial recognition and dendritic cell activation^{1,2}. TLRs 7 and 9 can recognize nucleic acids^{3–6} and trigger signalling cascades that activate plasmacytoid dendritic cells to produce interferon- α (IFN- α) (refs 7, 8). TLR7/9-mediated dendritic cell activation is critical for antiviral immunity but also contributes to the pathogenesis of systemic lupus erythematosus, a disease in which serum IFN- α levels are elevated owing to plasmacytoid dendritic cell activation^{8,9}. TLR7/9-induced IFN- α induction depends on a molecular complex that contains a TLR adaptor, MyD88, and IFN regulatory factor 7 (IRF-7) (refs 10–14), but the underlying molecular mechanisms are as yet unknown. Here we show that I κ B kinase- α (IKK- α) is critically involved in TLR7/9-induced IFN- α production. TLR7/9-induced IFN- α production was severely impaired in IKK- α -deficient plasmacytoid dendritic cells, whereas inflammatory cytokine induction was decreased but still occurred. Kinase-deficient IKK- α inhibited the ability of MyD88 to activate the *Ifna* promoter in synergy with IRF-7. Furthermore, IKK- α associated with and phosphorylated IRF-7. Our results identify a role for IKK- α in TLR7/9 signalling, and highlight IKK- α as a potential target for manipulating TLR-induced IFN- α production.

Among the TLRs, TLR9 subfamily members are characteristic in their ability to induce type-I interferons (such as IFN- α) from plasmacytoid dendritic cells (PDCs)⁷. Both TLR7 and TLR9 associate with a cytoplasmic adaptor, MyD88, that is essential for the induction of both inflammatory cytokine and type-I interferon production. The transcription factor IRF-7 and the serine/threonine kinase IRAK-1 (interleukin-1 receptor-associated kinase 1) are critical for TLR7/9-induced type-I interferon but not cytokine production, indicating bifurcation of the signalling pathways downstream of MyD88^{12,13}. IRF-7 associates with MyD88, tumour-necrosis factor (TNF) receptor-associated factor 6 (TRAF6) and IRAK-1, and this association is critical for induction of *Ifna* gene expression^{10–12}. However, the detailed molecular mechanisms of TLR7/9-induced IFN- α production have not yet been identified.

The IKK family contains four members, all of which are involved in a variety of immune responses^{15,16}. IKK- α and IKK- β are serine/threonine kinases that were initially isolated based on their ability to phosphorylate I κ B proteins and to activate NF- κ B. IKK- β was found to be more critical than IKK- α for both TLR- and cytokine-induced NF- κ B activation leading to inflammatory cytokine production¹⁶. The other two IKKs, TANK-binding kinase 1 (TBK-1) and IKK- ϵ (also called as IKK- ι), are not involved in NF- κ B activation but have critical roles in TLR3/4-induced *Ifnb* gene expression through IRF-3 phosphorylation^{17,18}. IKK- α , although dispensable for the canonical

NF- κ B activating pathway, is essential for a noncanonical pathway downstream of LT- β R or BAFF-R (B cell-activation factor belonging to the TNF family receptor)¹⁶. IKK- α is important for keratinocyte differentiation as well as for B-cell maturation and Peyer's patch formation, but only some of these require IKK- α kinase activity^{19–24}. However, it remains unknown whether or how IKK- α is implicated in TLR signalling.

IKK- α -deficient mice die soon after birth because of impaired keratinocyte differentiation^{19–21}. We therefore established bone

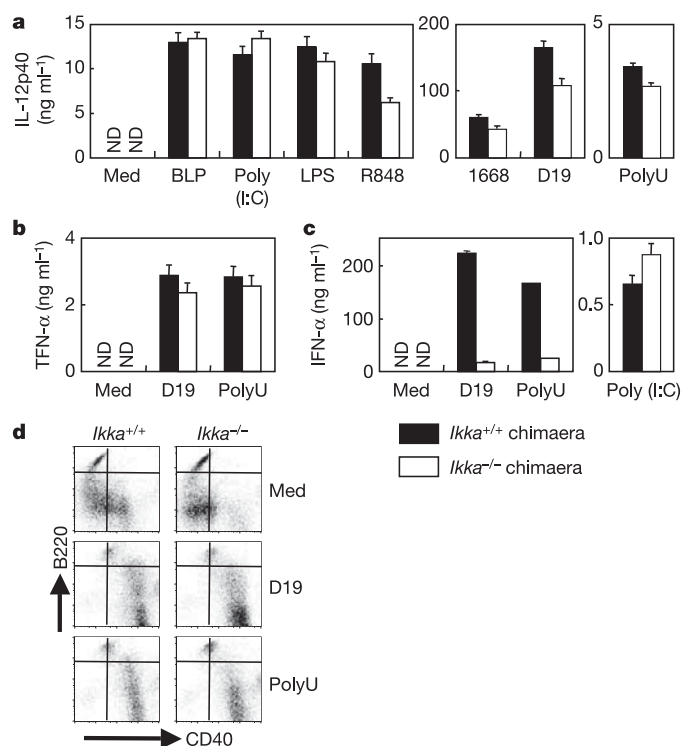


Figure 1 | Role for IKK- α in IFN- α production from TLR7/9-stimulated BMDCs *in vitro*. **a–d**, Flt3L-induced BMDCs from *I\kappa\kappa\alpha^{+/+}* or *I\kappa\kappa\alpha^{-/-}* chimaeras were unstimulated (med) or stimulated with 100 ng ml⁻¹ BLP, 50 μ g ml⁻¹ poly(I:C), 100 ng ml⁻¹ LPS, 100 nM R848, 1 μ M ODN1668, 3 μ M D19 or 25 μ g ml⁻¹ polyU in complex with lipofectamine 2000 for 20–24 h. In **c**, poly(I:C) was complexed with lipofectamine 2000 and used at 25 μ g ml⁻¹. Cytokine levels (**a–c**) were analysed by ELISA. Error bars show s.d. ND indicates less than 100 pg ml⁻¹. **d**, FACS profile of Flt3L-induced BMDCs.

¹Laboratory for Host Defense and ²Laboratory for Immunogenomics, RIKEN Research Center for Allergy and Immunology, Suehiro-cho 1-7-22, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. ³Division of Molecular Immunology, Institute for Enzyme Research, University of Tokushima, Kuramoto-cho 3-18-15, Tokushima, Tokushima 770-8503, Japan.

⁴Department of Host Defense, Research Institute for Microbial Diseases and ERATO, Japan Science and Technology Agency, Yamadaoka 3-1, Suita, Osaka 565-0871, Japan.

⁵Kazusa DNA Research Institute, Kazusa-Kamatari 2-6-7, Kisarazu, Chiba 292-0818, Japan.

marrow chimaeric mice and generated Flt3L-induced bone-marrow-derived dendritic cells (BMDCs). Both PDCs ($CD11c^{+}B220^{+}$) and conventional dendritic cells (cDCs, $CD11c^{+}B220^{-}$) were generated from *Ikka* knockout (*Ikka*^{-/-}) Flt3L-induced BMDCs (data not shown; *Ikka* is also known as *Chuk*). Both wild-type and *Ikka*^{-/-} Flt3L-induced BMDCs increased production of the cytokine interleukin (IL)-12p40 in response to bacterial lipopeptide (BLP, a TLR2 ligand), poly(I:C) (a TLR3 ligand), lipopolysaccharide (LPS, a TLR4 ligand), R848 (a TLR7 ligand) and ODN1668 (a TLR9 ligand), although IL-12p40 induction by TLR7/9 signalling was slightly lower in *Ikka*^{-/-} BMDCs (Fig. 1a). TNF- α induction was comparable between wild-type and *Ikka*^{-/-} BMDCs (Fig. 1b). Furthermore, ligands for TLR2, TLR3, TLR4, TLR7 and TLR9 could enhance cell-surface expression of co-stimulatory molecules such as CD40 and CD86 both in wild-type and *Ikka*^{-/-} BMDCs (Fig. 1d and data not shown). Thus, these functions common to TLR family members are not dependent on IKK- α .

D/A-type CpG DNA (D19)²⁵⁻²⁷ and polyuridylic acid (polyU)^{4,5} can induce robust production of IFN- α from PDCs in a TLR9- and TLR7-dependent manner, respectively. IFN- α production by these stimuli was severely impaired in *Ikka*^{-/-} BMDCs (Fig. 1c). However, poly(I:C)-induced IFN- α production was not impaired in *Ikka*^{-/-} BMDCs (Fig. 1c). Thus, IKK- α is involved in TLR7/9-mediated IFN- α induction.

In wild-type mice, PDCs are present in the bone marrow and

spleen⁸. PDC development was normal in *Ikka*^{-/-} chimaeras (Supplementary Fig. 1a, c). However, TLR7/9-induced IFN- α production was severely decreased in *Ikka*^{-/-} chimaeras, and IL-12p40 induction was mildly impaired (Supplementary Fig. 1b, d).

Flt3L-induced BMDCs and splenic $CD11c^{+}$ cells contain both PDCs and cDCs. TLR7/9-induced IFN- α production depends on PDCs, but IL-12p40 and TNF- α production is derived from both PDCs and cDCs. We therefore analysed cytokine production from PDCs and cDCs separately. IFN- α production from TLR7/9-stimulated *Ikka*^{-/-} PDCs was severely impaired. *Ikka*^{-/-} PDCs and cDCs produced comparable or slightly decreased amounts of IL-12p40 and TNF- α compared to wild-type PDCs and cDCs (Fig. 2a–c and Supplementary Fig. 2). Responses of purified *Ikka*^{-/-} PDCs to different doses of TLR7/9 agonists also showed severe defects in the expression of several *Ifna* genes (Fig. 2d and Supplementary Fig. 3). *Ifnb* expression was reduced in *Ikka*^{-/-} PDCs, but not as prominently as *Ifna* expression. IFN- α production was also severely impaired in *Ikka*^{-/-} PDCs infected with vesicular stomatitis virus (VSV), which can induce IFN- α in a TLR7-dependent manner (Fig. 2e). Injecting mice with TLR7 agonists can increase serum IFN- α levels in a PDC-dependent manner, but this effect was severely impaired in *Ikka*^{-/-} chimaeras (Fig. 2f). These results show that IKK- α is essential for IFN- α production in response to TLR7/9 signalling by PDCs *in vivo*.

Infection with Newcastle disease virus (NDV) can induce mouse

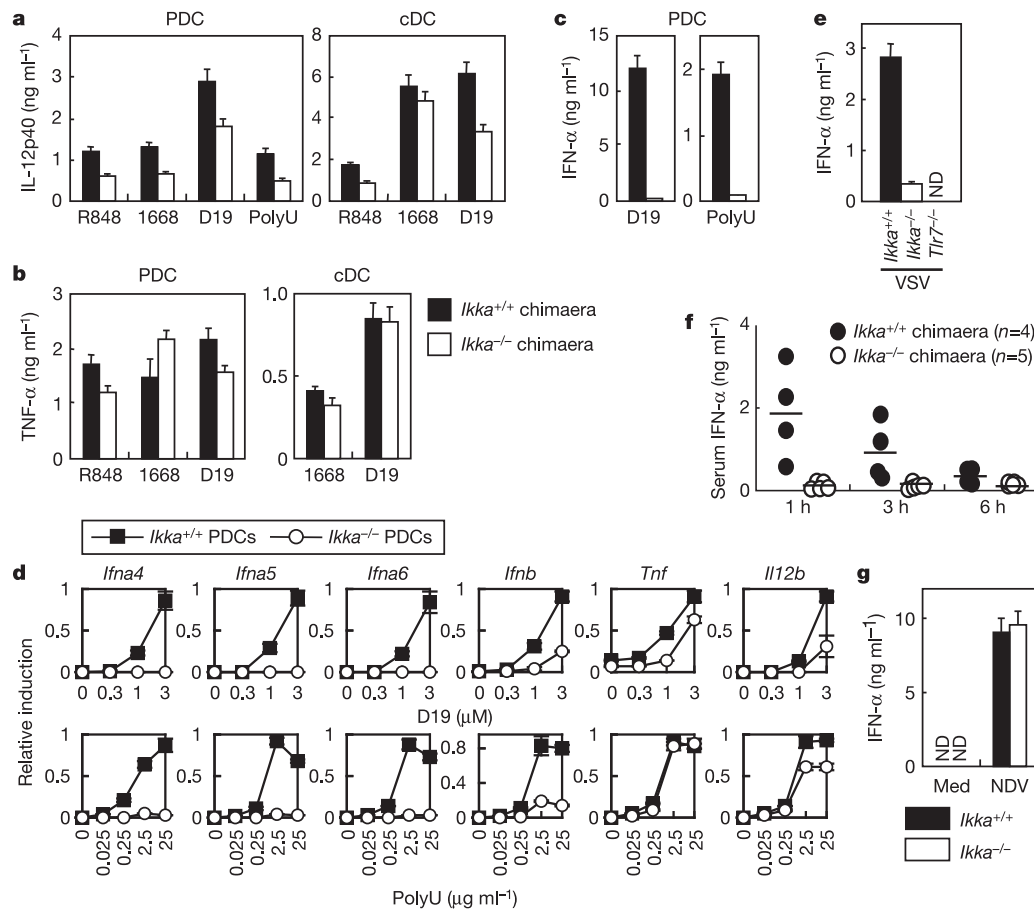


Figure 2 | Impairment of TLR7/9-induced IFN- α production in *Ikka*^{-/-} PDCs. **a–c**, Flt3L-induced bone-marrow-derived PDCs and cDCs were stimulated with TLR7 agonists (R848, PolyU) or TLR9 agonists (ODN1668, D19) (concentrations same as Fig. 1). Levels of IL-12p40 (**a**), TNF- α (**b**) and IFN- α (**c**) in cell-culture supernatants are shown. **d**, Expression of IFN- α / β and inflammatory cytokine messenger RNAs in *Ikka*^{+/+} and *Ikka*^{-/-} Flt3L-induced PDCs. PDCs were stimulated for 3 h with indicated

concentrations of D19 or PolyU, then subjected to quantitative real-time RT-PCR analysis. Data were normalized to the level of 18S rRNA expression. **e**, IFN- α induction in VSV-infected, Flt3L-induced PDCs. **f**, Serum IFN- α levels in R848-injected mice. Bars indicate means. **g**, IFN- α induction in NDV-infected *Ikka*^{+/+} and *Ikka*^{-/-} MEFs. Error bars show s.d. ND indicates less than 100 pg ml⁻¹.

embryonic fibroblasts (MEFs) to produce IFN- α in a manner dependent on the cytoplasmic RNA helicase RIG-I (refs 13, 28). We next examined whether IKK- α is involved in this cytosolic pathway leading to IFN- α induction. Upon NDV infection, *Ifna* expression and production were enhanced in both wild-type and *Ikk α ^{-/-}* MEFs (Fig. 2g and Supplementary Fig. 4). IKK- α is therefore dispensable for this cytosolic pathway-induced production of IFN- α in MEFs. PDCs have both a TLR-dependent and a cytosolic pathway for virus recognition²⁹. It remains unclear whether or how IKK- α is involved in the cytosolic pathway in PDCs.

To further elucidate how IKK- α is involved in IFN- α induction, we analysed the effects of IKK- α on the *Ifna* promoter, which is mediated by IRF-7. Expression of the TLR adaptor MyD88 alone failed to activate the *Ifna* promoter, but did increase IRF-7-induced promoter activation (Fig. 3a). Expression of kinase-deficient IKK- α

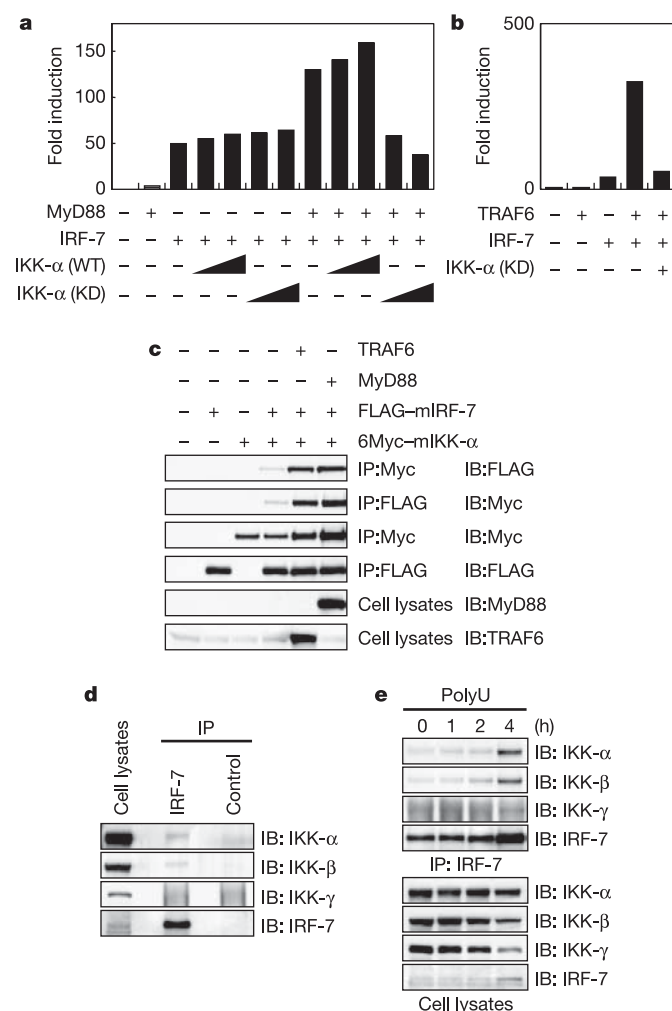


Figure 3 | Role for IKK- α in IRF-7-mediated signalling. **a, b**, 293T cells were transiently transfected with an *Irfn4* promoter-driven luciferase reporter plasmid alone or together with a combination of expression vectors for MyD88 (120 ng), TRAF6 (120 ng), IRF-7 (12 ng), wild-type IKK- α (WT, at 240 or 600 ng) or kinase-deficient IKK- α (KD, at 240 or 600 ng). In **b**, 240 ng of the kinase-deficient IKK- α expression vector was used. After 24 h, cell lysates were prepared and promoter activity was measured by luciferase assay. **c**, Interaction between IKK- α and IRF-7 in 293T cells. IP, immunoprecipitation; IB, immunoblot. **d**, Unstimulated Flt3L-induced BMDCs were lysed, immunoprecipitated with control or anti-IRF-7 antibodies and immunoblotted with the indicated antibodies. **e**, IRF-7 immunoprecipitates were prepared from PolyU-stimulated Flt3L-induced BMDCs and analysed by immunoblots. Results are representative of three separate experiments.

did not affect IRF-7-mediated promoter activation, but inhibited the enhancing effects of MyD88 (Fig. 3a). Expression of wild-type IKK- α slightly increased *Ifna* promoter activation by MyD88 and IRF-7. Similar to MyD88, another adaptor molecule, TRAF6, could also activate the *Ifna* promoter in synergy with IRF-7 (Fig. 3b). This synergistic effect was also inhibited by expression of kinase-deficient IKK- α . Thus, IKK- α kinase activity is required for MyD88 or TRAF6 to activate *Ifna* promoter in synergy with IRF-7.

We next tested whether IKK- α associates with IRF-7. In 293T cells, IKK- α co-immunoprecipitated with IRF-7 (Fig. 3c), and this association was significantly enhanced by coexpression of MyD88 or TRAF6 (Fig. 3c). Furthermore, in Flt3L-induced BMDCs and IRF-7-expressing BMDCs induced with granulocyte-macrophage colony-stimulating factor (GM-CSF), IKK- α was detected in IRF-7 immunoprecipitates, but not in control immunoprecipitates (Fig. 3d and Supplementary Fig. 5a). This association was observed before stimulation; after stimulation with TLR7/9 agonists, the amount of co-immunoprecipitated IKK- α increased, and was accompanied by an increase in IRF-7 protein levels (Fig. 3e and Supplementary Fig. 5b). IKK- β was also co-immunoprecipitated in anti-IRF-7 immunoprecipitates, and the co-immunoprecipitated

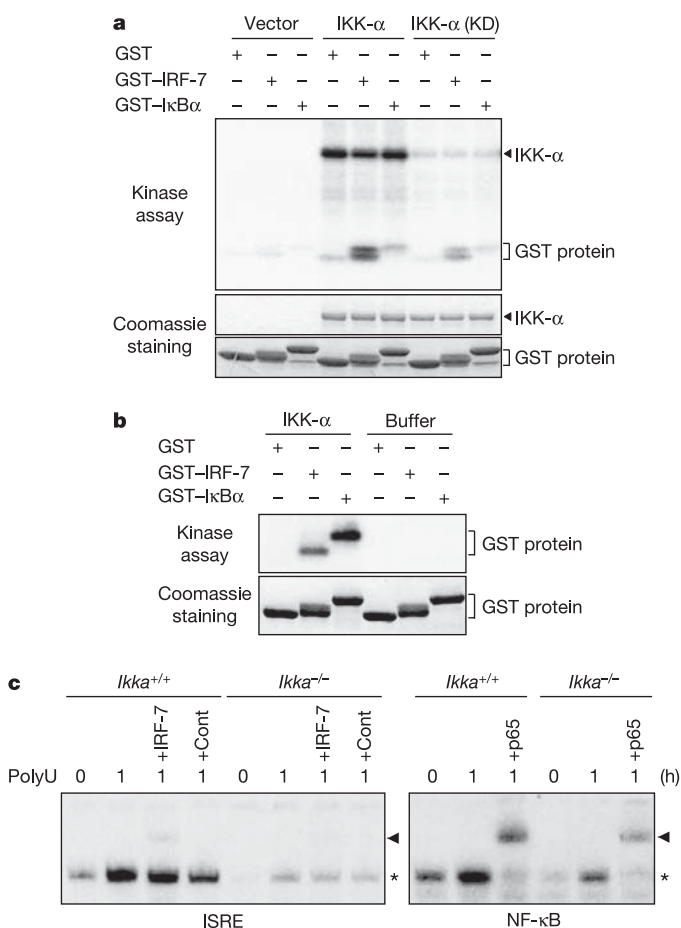


Figure 4 | IRF-7 phosphorylation and activation by IKK- α . **a**, 293T cells were transfected with a control vector or expression vectors for wild-type IKK- α or kinase-deficient IKK- α (KD). Immunoprecipitations were performed with an anti-Myc-tag monoclonal antibody, and immunoprecipitates were subjected to *in vitro* kinase assay. **b**, Recombinant IKK- α was subjected to an *in vitro* kinase assay. **c**, Flt3L-induced PDCs were unstimulated or stimulated with polyU for 1 h and nuclear extracts were prepared. Then ISRE and NF- κ B binding activity (asterisks) was determined by EMSA. Supershifted bands are indicated with arrowheads. Anti-IRF-7, rabbit Ig (Cont) and anti-p65 antibodies were used for supershift analysis.

amount increased after stimulation; in contrast, IKK- γ was hardly detected. Thus, IKK- β as well as IKK- α are present in the IRF-7-containing molecular complex^{10,11}. Further analysis of IKK- β - and IKK- γ -deficient mice is necessary to clarify whether IKK- β or IKK- γ is critical for TLR7/9-induced IFN- α production.

NF- κ B-inducing kinase (NIK) interacts with and phosphorylates IKK- α ^{15,16}. Both NIK and IKK- α are critical for the noncanonical NF- κ B activation pathway. We next analysed *aly/aly* mice, which carry a nonfunctional NIK mutant. TLR7/9-induced IFN- α and IL-12p40 production was normal in *aly/aly* splenic cells (Supplementary Fig. 6a, b) and Flt3L-induced BMDCs (Supplementary Fig. 6c, d), indicating that IFN- α induction in TLR7/9 signalling is independent of NIK.

IRF-7 resides in the cytoplasm in an inactive form. Upon stimulation, IRF-7 is phosphorylated and translocates into the nucleus, where it has a role in activating *Ifna* gene expression¹³. We therefore investigated whether IKK- α can phosphorylate IRF-7. Immunoprecipitates of wild-type IKK- α phosphorylated IRF-7 and I κ B α (Fig. 4a), but this phosphorylating activity was strongly reduced with kinase-deficient IKK- α . Any remaining phosphorylation in immunoprecipitates of kinase-deficient IKK- α is likely to be due to co-precipitation of small amounts of endogenous IKK- α or other kinase(s) (Fig. 4a). Baculovirus-expressed IKK- α also phosphorylated IRF-7 (Fig. 4b). Notably, IFN-stimulated response element (ISRE)-binding activity of IRF-7 was impaired in polyU-stimulated *Ikka*^{-/-} PDCs, but NF- κ B binding activity was normal in these cells, indicating defective IRF-7 activation in the absence of IKK- α (Fig. 4c).

TBK-1 and IKK- ϵ are also reported to phosphorylate IRF-7 and to activate the *Ifna* promoter³⁰. However, TLR9-induced IFN- α production is not impaired in TBK-1- or IKK- ϵ -deficient mice, indicating that either kinase is dispensable for TLR9-induced IFN- α production¹⁰. Furthermore, our results with *Ikka*^{-/-} mice clearly show that neither TBK-1 nor IKK- ϵ can substitute for IKK- α in terms of TLR7/9-induced IFN- α production.

The most severe defect we observed in IKK- α -deficient DCs was in IFN- α induction. Notably, *Ifnb* expression was not as severely impaired as *Ifna* induction in IKK- α deficiency, whereas induction of both genes was found to be almost completely abolished in IRF-7-deficient PDCs¹³. *Ifna* expression is strongly dependent on IRF-7, but NF- κ B also contributes to *Ifnb* expression together with IRF-7. In the absence of IKK- α , IRF-7 can be activated by other kinases. Such residual IRF-7 activity, although not detectable in electrophoretic mobility shift assays (EMSA) of ISRE (Fig. 4c), can still lead to some degree of *Ifnb* expression together with intact NF- κ B activity. Our present findings clearly support the proposal that IKK- α is a critical component of the cytoplasmic transductional-transcriptional processor¹¹ leading to induction of *Ifna* expression. It is unclear at present how IKK- α cooperates with IRAK-1. We suggest that IKK- α is a potential molecular target not only for manipulating antiviral immunity but also for treating autoimmune disorders such as systemic lupus erythematosus, in which IFN- α production is elevated.

METHODS

Mice. The generation of *Ikka*^{-/-} mice has been previously described¹⁹. The *+/-aly* or *aly/aly* mice were purchased from CLEA Japan. All mice were maintained under specific pathogen-free conditions in the animal facility of the RIKEN Research Center for Allergy and Immunology and used after backcrossing with C57BL/6 mice at least seven times. Bone marrow chimaeric mice were generated by transferring *Ikka*^{+/+} or *Ikka*^{-/-} fetal liver cells from embryos into C57BL/6 (CD45.1) mice as described previously²².

Cells. Flt3L-induced BMDCs and splenic CD11c⁺ cells were generated as described previously²⁷. PDCs (CD11c⁺B220⁺) and cDCs (CD11c⁺B220⁻) were sorted using a FACS Vantage flow cytometer (BD Biosciences).

Measurement of cytokine production. Cells were stimulated for 20–24 h with the indicated stimuli. Concentrations of IL-12p40, TNF- α and IFN- α in cell-culture supernatants were measured by enzyme-linked immunosorbent assay

(ELISA) as described previously²⁷. Serum IFN- α levels were also measured in mice injected intravenously with 50 nmol R848.

Reverse transcriptase-polymerase chain reaction (RT-PCR) analysis. Quantitative real-time RT-PCR analysis was performed using a 7000 Sequence detector (Applied Biosystems). Reactions were performed using primers against the following sequences: 18S rRNA (internal control), *Ifna4*, *Ifna5*, *Ifna6*, *Ifnb*, *Tnf* and *Il12b* (TaqMan Gene Expression Assays, Applied Biosystems).

Electrophoretic mobility shift assays (EMSA). EMSA were performed as described previously¹⁷. ³²P-labelled double-stranded oligonucleotides corresponding to the ISRE of the human *ISG15* gene (also known as *GIP2*; 5'-CTCGGAAAGGGAAACCGAACTGAAGCC-3') and the NF- κ B binding site (5'-ATCAGGGACTTTCGGCTGGGGACTTTCGCC-3') were used. For supershift analysis, anti-IRF-7 (Abcam) or anti-p65 (Santa Cruz) antibodies were incubated with nuclear extracts before the addition of probes. Rabbit IgG (Santa Cruz) was used as a control antibody.

Received 13 December 2005; accepted 9 February 2006.

- Medzhitov, R. Toll-like receptors and innate immunity. *Nature Rev. Immunol.* **1**, 135–145 (2001).
- Takeda, K., Kaisho, T. & Akira, S. Toll-like receptors. *Annu. Rev. Immunol.* **21**, 335–376 (2003).
- Hemmi, H. et al. A Toll-like receptor recognizes bacterial DNA. *Nature* **408**, 740–745 (2000).
- Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S. & Reis, E. S. C. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **303**, 1529–1531 (2004).
- Heil, F. et al. Species-specific recognition of single-stranded RNA via Toll-like receptor 7 and 8. *Science* **303**, 1526–1529 (2004).
- Lund, J. M. et al. Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc. Natl Acad. Sci. USA* **101**, 5598–5603 (2004).
- Wagner, H. The immunobiology of the TLR9 subfamily. *Trends Immunol.* **25**, 381–386 (2004).
- Colonna, M., Trinchieri, G. & Liu, Y. J. Plasmacytoid dendritic cells in immunity. *Nature Immunol.* **5**, 1219–1226 (2004).
- Blanco, P., Palucka, A. K., Gill, M., Pascual, V. & Banchereau, J. Induction of dendritic cell differentiation by IFN- α in systemic lupus erythematosus. *Science* **294**, 1540–1543 (2001).
- Kawai, T. et al. Interferon- α induction through Toll-like receptors involves a direct interaction of IRF7 with MyD88 and TRAF6. *Nature Immunol.* **5**, 1061–1068 (2004).
- Honda, K. et al. Role of a transductional-transcriptional processor complex involving MyD88 and IRF-7 in Toll-like receptor signaling. *Proc. Natl Acad. Sci. USA* **101**, 15416–15421 (2004).
- Uematsu, S. et al. Interleukin-1 receptor-associated kinase-1 plays an essential role for Toll-like receptor (TLR)7- and TLR9-mediated interferon- α induction. *J. Exp. Med.* **201**, 915–923 (2005).
- Honda, K. et al. IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature* **434**, 772–777 (2005).
- Honda, K. et al. Spatiotemporal regulation of MyD88–IRF-7 signalling for robust type-I interferon induction. *Nature* **434**, 1035–1040 (2005).
- Hayden, M. S. & Ghosh, S. Signaling to NF- κ B. *Genes Dev.* **18**, 2195–2224 (2004).
- Bonizzi, G. & Karin, M. The two NF- κ B activation pathways and their role in innate and adaptive immunity. *Trends Immunol.* **25**, 280–288 (2004).
- Hemmi, H. et al. The roles of two I κ B kinase-related kinases in lipopolysaccharide and double stranded RNA signaling and viral infection. *J. Exp. Med.* **199**, 1641–1650 (2004).
- Perry, A. K., Chow, E. K., Goodnough, J. B., Yeh, W. C. & Cheng, G. Differential requirement for TANK-binding kinase-1 in type I interferon responses to Toll-like receptor activation and viral infection. *J. Exp. Med.* **199**, 1651–1658 (2004).
- Takeda, K. et al. Limb and skin abnormalities in mice lacking IKK α . *Science* **284**, 313–316 (1999).
- Hu, Y. et al. Abnormal morphogenesis but intact IKK activation in mice lacking the IKK α subunit of I κ B kinase. *Science* **284**, 316–320 (1999).
- Li, Q. et al. IKK1-deficient mice exhibit abnormal development of skin and skeleton. *Genes Dev.* **13**, 1322–1328 (1999).
- Kaisho, T. et al. I κ B kinase α is essential for mature B cell development and function. *J. Exp. Med.* **193**, 417–426 (2001).
- Senftleben, U. et al. Activation by IKK α of a second, evolutionary conserved, NF- κ B signaling pathway. *Science* **293**, 1495–1499 (2001).
- Matsushima, A. et al. Essential role of nuclear factor (NF)- κ B-inducing kinase and inhibitor of κ B (I κ B) kinase α in NF- κ B activation through lymphotoxin β receptor, but not through tumor necrosis factor receptor I. *J. Exp. Med.* **193**, 631–636 (2001).
- Verthelyi, D., Ishii, K. J., Gursel, M., Takeshita, F. & Klinman, D. M. Human peripheral blood cells differentially recognize and respond to two distinct CPG motifs. *J. Immunol.* **166**, 2372–2377 (2001).
- Krug, A. et al. Identification of CpG oligonucleotide sequences with high induction of IFN- α / β in plasmacytoid dendritic cells. *Eur. J. Immunol.* **31**, 2154–2163 (2001).

27. Hemmi, H., Kaisho, T., Takeda, K. & Akira, S. The roles of Toll-like receptor 9, MyD88, and DNA-dependent protein kinase catalytic subunit in the effects of two distinct CpG DNAs on dendritic cell subsets. *J. Immunol.* **170**, 3059–3064 (2003).
28. Kato, H. *et al.* Cell type-specific involvement of RIG-I in antiviral response. *Immunity* **23**, 19–28 (2005).
29. Hochrein, H. *et al.* Herpes simplex virus type-1 induces IFN- α production via Toll-like receptor 9-dependent and -independent pathways. *Proc. Natl Acad. Sci. USA* **101**, 11416–11421 (2004).
30. Sharma, S. *et al.* Triggering the interferon antiviral response through an IKK-related pathway. *Science* **300**, 1148–1151 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank N. Iwami, Y. Fukuda, N. Okita and E. Haga for

technical assistance; Y. Hachiman and H. Fujimoto for cell sorting; and S. Haraguchi and T. Oyanagi for secretarial assistance. We also thank R. Triendl for critical review of the manuscript. This work was supported by the Advanced and Innovational Research Program in Life Sciences, the Special Coordination Fund for Promoting Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and Grant-in-Aid for Scientific Research of MEXT and the Japan Society for the Promotion of Science, the Uehara Memorial Foundation, the Naito Foundation, the Novartis Foundation for the Promotion of Science, and RIKEN Strategic programs. T.S. is supported by a RIKEN Junior Research Associate grant.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.K. (tkaisho@rcai.riken.jp).

LETTERS

The reversibility of mitotic exit in vertebrate cells

Tamara A. Potapova¹, John R. Daum¹, Bradley D. Pittman¹, Joanna R. Hudson¹, Tara N. Jones¹, David L. Satinover², P. Todd Stukenberg² & Gary J. Gorsky¹

A guiding hypothesis for cell-cycle regulation asserts that regulated proteolysis constrains the directionality of certain cell-cycle transitions^{1,2}. Here we test this hypothesis for mitotic exit, which is regulated by degradation of the cyclin-dependent kinase 1 (Cdk1) activator, cyclin B^{3–5}. Application of chemical Cdk1 inhibitors to cells in mitosis induces cytokinesis and other normal aspects of mitotic exit, including cyclin B degradation. However, chromatid segregation fails, resulting in entrapment of chromatin in the midbody. If cyclin B degradation is blocked with a proteasome inhibitor or by expression of non-degradable cyclin B, Cdk1 inhibitors will nonetheless induce mitotic exit and cytokinesis. However, if after mitotic exit, the Cdk1 inhibitor is washed free from cells in which cyclin B degradation is blocked, the cells can revert back to M phase. This reversal is characterized by chromosome recondensation, nuclear envelope breakdown, assembly of microtubules into a mitotic spindle, and in most cases, dissolution of the midbody, reopening of the cleavage furrow, and realignment of chromosomes at the metaphase plate. These findings demonstrate that proteasome-dependent degradation of cyclin B provides directionality for the M phase to G1 transition.

Cdk1, the major regulator of mitotic progression, is activated through binding of cyclin A or B. Cyclin A is degraded during prometaphase when chromosomes move to align at the metaphase plate^{6,7}. Cyclin B degradation begins at metaphase and continues during chromatid segregation in anaphase and exit from M phase⁵. Cytokinesis is initiated shortly after anaphase onset. Cdk1 inactivation and dephosphorylation of Cdk1 substrates during mitotic exit probably serve as timing mechanisms to ensure that cytokinesis occurs after chromatid separation^{8–12}. For example, high Cdk1 activity before anaphase blocks the accumulation of the cytokinetic regulators aurora B and MKLP1 at the cleavage furrow and on the microtubules of the spindle midzone^{13–15}.

Flavopiridol is a potent inhibitor of Cdk1¹⁶. We found that treatment of vertebrate cells in mitosis with flavopiridol resulted in premature mitotic exit accompanied by cytokinesis (Fig. 1a and Supplementary Video 1). Similar results have recently been found for the Cdk inhibitor BMI-1026 (ref. 17). Flavopiridol induced the microtubule network to undergo changes characteristic of anaphase and mitotic exit. The spindle poles moved apart, and microtubule bundles formed in the spindle midzone and at the equatorial cortex. Even though chromatid separation did not occur, cytokinetic furrows formed and ingressed to completion. The cleavage furrow trapped chromosomes in the midbody, resulting in a 'cut' phenotype. Nevertheless, the chromosomes decondensed and nuclear envelopes reformed. Eventually, cytoplasmic contractile activity diminished as cells flattened fully onto the substratum, and the microtubule array established an interphase pattern.

During normal mitotic exit, Cdk1 activity is reduced by ubiquitylation and proteasome-mediated degradation of cyclin B^{3,5}. Proteasome inhibitors such as MG132 induce mitotic cells to arrest at

metaphase. We found that treatment with flavopiridol overrides metaphase arrest induced with MG132, causing mitotic exit and cytokinesis accompanied by chromosome decondensation and reformation of the nuclear envelope (Fig. 1b and Supplementary Video 2). The proteolysis of cyclin B at mitotic exit is thought to ensure the unidirectionality of the M phase to G1 transition². In cells with proteasome inhibition, we found that flavopiridol-induced mitotic exit was reversible. Upon flavopiridol removal, cells that had exited mitosis could return to metaphase (Fig. 1c and Supplementary Videos 3, 4). The microtubules, having assumed an interphase configuration after flavopiridol-induced mitotic exit, reassembled a mitotic spindle when flavopiridol was removed. The midbody disappeared and the cytokinetic furrow retracted, the newly formed nuclear envelope dissolved, and the chromosomes recondensed, attached to spindle microtubules and realigned at the metaphase plate. These findings are summarized in Supplementary Fig. 1. We found that cells induced to reverse back to metaphase could subsequently undergo a second, normal mitotic exit (including chromatid separation and movement) as well as a second cytokinesis if the proteasome inhibitor was subsequently washed away (Supplementary Video 5). We used flavopiridol for most of our experiments because of its high potency as a Cdk1 inhibitor (Supplementary Fig. 2). However, some other chemical Cdk inhibitors also resulted in similar phenotypes (see Supplementary Information).

In *Xenopus* S3 cells arrested at metaphase with MG132, aurora B kinase was concentrated at centromeres and MKLP1 was localized diffusely in the cytoplasm (Fig. 1d). Upon treatment with flavopiridol, aurora B and MKLP1 rapidly accumulated at the equatorial cell cortex, associating with the nascent cleavage furrows and the midzone overlying the microtubule bundles, and eventually becoming highly concentrated at the midbody. These translocation events were consistent with the typical relocation of aurora B and MKLP1 in normal anaphase, and were also reversible: in cells that were induced through mitotic exit and then allowed to revert back to metaphase, aurora B and MKLP1 returned to their typical metaphase distributions (Fig. 1d, 60 min).

Metaphase cells treated with flavopiridol in the presence of proteasome inhibitor advanced through mitotic exit and cytokinesis, reaching the midbody stage with decondensed chromosomes and reassembled nuclear envelopes 25 min after treatment (Fig. 1b). The reversibility of mitotic exit was dependent on the duration of exposure to flavopiridol. Treatment with 5 μ M for 17–30 min resulted in most cells (81 out of 106) returning to metaphase upon flavopiridol washout (Table 1). With longer exposure to flavopiridol, the proportion of cells that underwent reversal declined. However, a few cells (3 out of 24) still reversed when flavopiridol was removed 80 min after its addition. We detected no reversal among cells treated for 90 min ($n = 26$). Reversibility was entirely dependent on the presence of the proteasome inhibitor. In medium without

¹Program in Molecular, Cell and Developmental Biology, Oklahoma Medical Research Foundation, 825 NE 13th Street, Oklahoma City, Oklahoma 73104, USA. ²Department of Biochemistry and Molecular Genetics, University of Virginia Medical School, 1300 Jefferson Park Avenue, Charlottesville, Virginia 22908, USA.

proteasome inhibitor, premature mitotic exit and cytokinesis induced with flavopiridol did not reverse in any of 13 prometaphase cells when the flavopiridol was removed at 17–20 min.

Previous studies have shown that several early mitotic regulators (for example, cyclin A) are normally degraded during prophase and prometaphase, as well as during the prometaphase arrest induced with microtubule drugs^{6,7}. We treated *Xenopus* S3 cells with the microtubule drug nocodazole to allow degradation of early mitotic regulators. Cells were then released from nocodazole to medium containing proteasome inhibitor. The cells assembled spindles and arrested at metaphase, but were fully capable of undergoing flavopiridol-induced mitotic exit and reversal (Fig. 2a and Supplementary Video 6). This result is consistent with the interpretation that mitotic exit reversal in cells treated with proteasome inhibitor is due to the protection of cyclin B from degradation, not due to the preservation of early mitotic regulators.

We next measured cell-cycle markers during induced mitotic exit and reversal. We treated nocodazole-arrested HeLa cells with flavopiridol according to the scheme depicted in Fig. 2b. We then calculated mitotic indices from the percentage of cells with condensed chromosomes (Fig. 2c, top), and immunoblotted samples for

various cell-cycle markers (Fig. 2c, bottom). Although the level of Cdk1 protein remained constant, flavopiridol induced mitotic exit and degradation of cyclins B1 and B2 in the absence of the proteasome inhibitor MG132 (Group 3, –MG132). In medium containing MG132, flavopiridol induced mitotic exit, but cyclin B1 and B2 were preserved (Group 3, +MG132). Flavopiridol removal at 25 min caused reversal back to M phase in the presence of MG132 (Group 5, +MG132, asterisk) but not in its absence (Group 5, –MG132). As expected, cyclin A, prominent when the initial population of mitotic cells was analysed (Group 1, –45 min), was substantially degraded during the subsequent 45 min incubation period before flavopiridol addition (Group 2). When flavopiridol was removed at 25 min in the mitotic exit reversal experiments, cyclin A levels were very low (Group 3) and thus unlikely to contribute significantly to the reactivation of Cdk1. After longer treatments with flavopiridol, some re-synthesis of cyclin A protein seemed to occur in cells treated with MG132 (Groups 4 and 5, +MG132). In summary, reversal of mitotic exit after flavopiridol removal correlates with the preservation of cyclins B1 and B2.

Flavopiridol was an effective inhibitor when purified Cdk1/cyclinB complex was used to phosphorylate histone H1 *in vitro*, showing

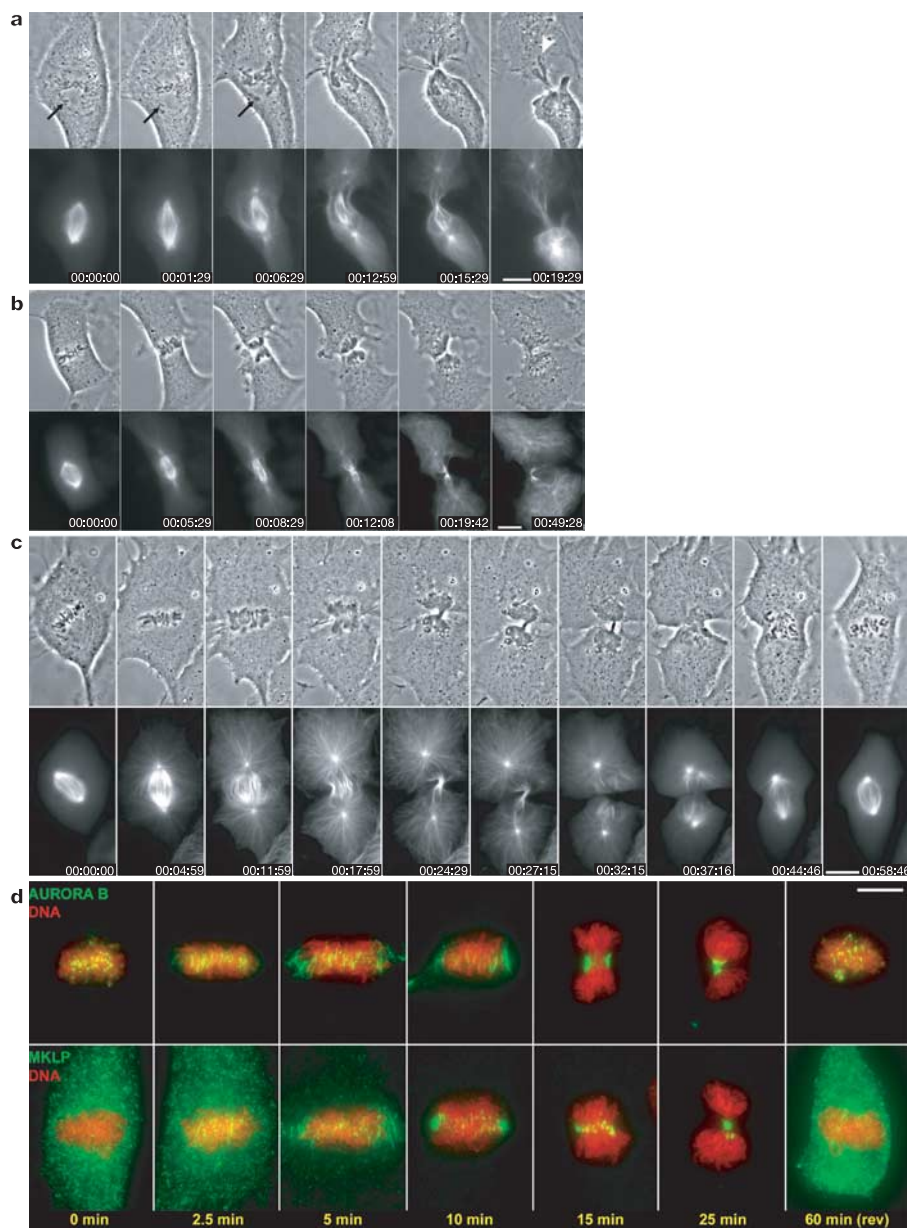


Figure 1 | The Cdk inhibitor flavopiridol induces reversible mitotic exit and cytokinesis if proteasome activity is inhibited. **a**, Treatment of mitotic cells with flavopiridol induces premature mitotic exit and cytokinesis without chromatid separation. A *Xenopus* S3 cell expressing α -tubulin–GFP was treated with 5 μ M flavopiridol at time 0. The black arrow indicates a chromosome not yet at metaphase when flavopiridol was added. The white arrowhead indicates the reassembled nuclear envelope. Supplementary Video 1 shows the complete video sequence. **b**, Flavopiridol induces mitotic exit in cells arrested at metaphase with MG132. An MG132-arrested cell was treated with 10 μ M flavopiridol at time 0. Supplementary Video 2 shows the complete video sequence. **c**, Flavopiridol-induced mitotic exit and cytokinesis are reversible if the proteasome is inhibited. An MG132-arrested cell was treated with 5 μ M flavopiridol at time 0, and flavopiridol was removed at 25 min. Supplementary Video 3 shows the complete video sequence. The top panels in **a–c** show phase-contrast images, and the bottom panels show α -tubulin–GFP fluorescence. Time in **a–c** is indicated as hours:minutes:seconds. **d**, Flavopiridol induces normal mitotic exit changes in the distributions of aurora B kinase and MKLP1. These are reversible upon flavopiridol removal. *Xenopus* S3 cells were incubated in medium containing MG132 and were fixed before and after treatment with 5 μ M flavopiridol. Samples were labelled for aurora B kinase (top panels) or MKLP1 (bottom panels). First and last images (0 min and 60 min) were scaled identically for brightness and contrast. Other images were scaled individually. Scale bars, 10 μ m.

Table 1 | Reversal of mitotic exit in *Xenopus* S3 cells upon removal of flavopiridol

Duration of flavopiridol treatment (min)	Number of cells imaged	Number of cells reversed through mitotic exit (% reversed)
17	29	28 (97)
20	27	25 (93)
25	27	17 (63)
30	23	11 (48)
33–37	17	8 (47)
40	22	6 (27)
50	19	4 (21)
60	17	1 (6)
70	24	1 (4)
80	24	3 (12)
90	26	0 (0)

higher potency than certain commonly used Cdk1 inhibitors such as olomoucine and roscovitine (Supplementary Fig. 2). In living cells, flavopiridol resulted in dephosphorylation of known Cdk1 substrates. In both the presence and absence of MG132, flavopiridol eliminated labelling with the TG-3 antibody that recognizes a Cdk1-catalysed phospho-epitope on nucleolin¹⁸ (Group 3). Cdh1, an activator of the mitotic ubiquitin ligase APC/C, is inhibited by Cdk1 phosphorylation during M phase¹⁹. Flavopiridol treatment resulted in an increase in the electrophoretic mobility of Cdh1, but only in the absence of MG132. Cdh1 dephosphorylation may be involved the ubiquitylation and breakdown of cyclin B after flavopiridol treatment in the absence MG132 (Group 3, –MG132).

Like *Xenopus* S3 cells, HeLa cells cultured in MG132 underwent a defined period after flavopiridol-induced mitotic exit when removal of flavopiridol led to reversal back to metaphase. When flavopiridol was removed 25 min after its addition, all cells (10 out of 10) reversed back to metaphase. When removed at 60 min, only 2 out of 12 cells (17%) underwent reversal.

To test whether preservation of cyclin B alone was sufficient for reversal of mitotic exit and cytokinesis, we performed a series of experiments using non-degradable cyclin B1 fused to green fluorescent protein (cyclinB1(R42A)–GFP) in place of the proteasome inhibitor. Most HeLa cells expressing non-degradable cyclin B arrested at metaphase, as has previously been shown¹². When these metaphase-arrested cells were treated with flavopiridol, they completed cytokinesis and exited mitosis. In some cells, flavopiridol treatment also induced chromatids to separate and to move to the poles before mitotic exit and cytokinesis. When flavopiridol was removed 25 min after its addition, 17 out of 37 transfected cells reverted back to M phase. In all 17 cells that underwent mitotic exit reversal, the chromosomes recondensed within 10 min of flavopiridol removal. Cells that had not separated their chromatids opened their cleavage furrows rapidly (average time, 13 min after flavopiridol removal) and reversed back to metaphase within 30 min (Fig. 3a and Supplementary Video 7). Flavopiridol addition induced chromatid separation in 7 of the 17 cells that subsequently underwent mitotic exit reversal. In these cells, the chromatids recondensed rapidly (within 10 min of flavopiridol removal) but opening of the cleavage furrow was somewhat delayed (average time 45 min after flavopiridol removal) (Fig. 3b and Supplementary Video 8). The separated chromatids remained near the centre of the cell but did not realign to form a metaphase plate. Two of the seven cells with separated chromatids did not show evidence of reopening their cleavage furrows at our final observation time (90 min after flavopiridol removal). Thus, cells with separated chromatids can undergo reversal of mitotic exit and cytokinesis, but predictably, the chromatids cannot realign at the metaphase plate if the sister chromatids have lost cohesion. Moreover, although cytokinesis is still reversible after chromatid segregation, reopening of the cleavage furrow generally takes longer to initiate. Trapped chromatin might physically facilitate re-opening of the furrow, or cells that separate their chromatids may have biochemical differences that restrict furrow

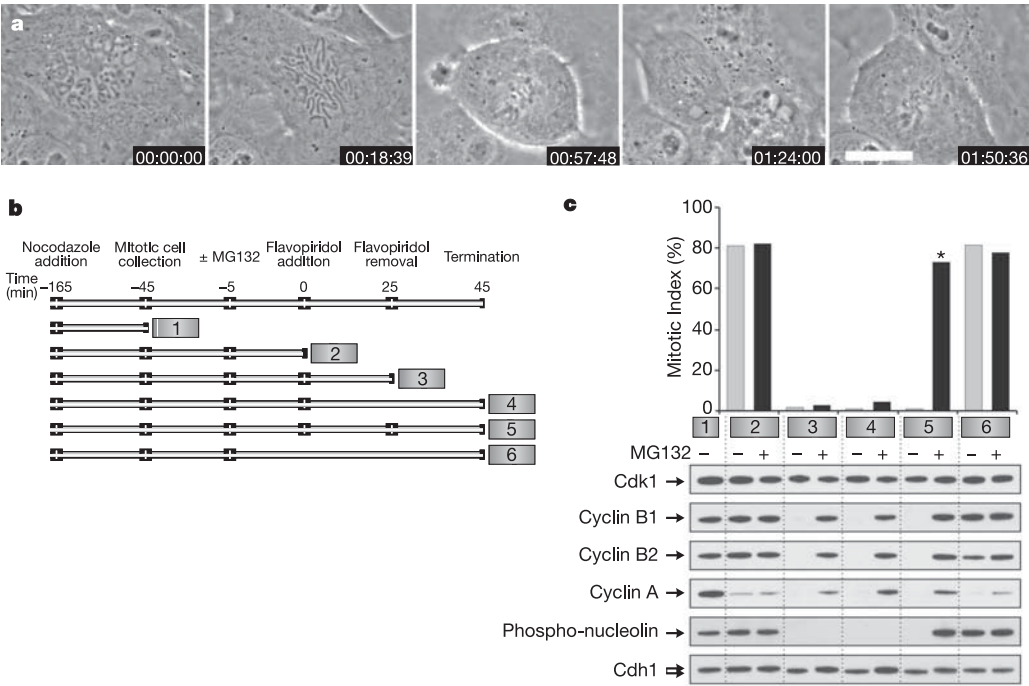


Figure 2 | Reversibility of mitotic exit requires preservation of cyclin B. **a**, A *Xenopus* S3 cell treated with 100 ng ml⁻¹ nocodazole at time 0. Nocodazole was removed 30 min after nuclear envelope breakdown, and the cell was arrested at metaphase with the addition of MG132. Flavopiridol at 5 μM was added at 1:04:00 and removed 25 min later, after mitotic exit. Supplementary Video 6 shows the complete video sequence. Scale bar,

10 μm. **b**, Experimental protocol used for producing samples from HeLa cells at various stages of induction and reversal of mitotic exit. **c**, Samples were analysed for mitotic index (the proportion of cells undergoing mitosis) and blotted for the indicated proteins. Asterisk indicates population that underwent mitotic exit reversal.

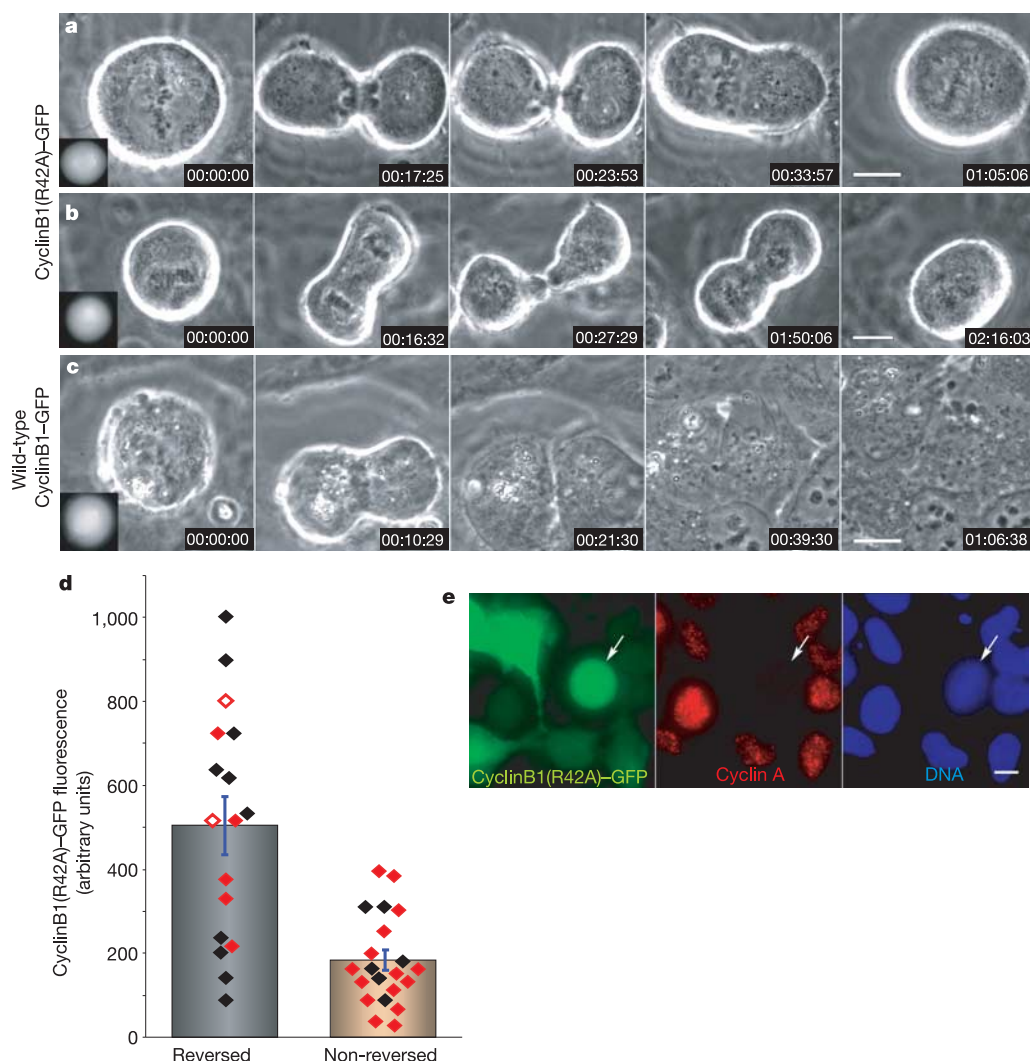


Figure 3 | Cells expressing non-degradable cyclin B1 undergo mitotic exit reversal. **a**, HeLa cells expressing non-degradable cyclin B1 can undergo mitotic exit reversal without segregating chromosomes. Flavopiridol was added at time 0 and removed at 25 min. Supplementary Video 7 shows the complete video sequence. **b**, HeLa cells expressing non-degradable cyclin B1 can undergo mitotic exit reversal after chromatid separation. Flavopiridol added and removed as in **a**. Supplementary Video 8 shows the complete video sequence. **c**, HeLa cells expressing wild-type cyclin B1 do not undergo reversal of mitotic exit. Flavopiridol added and removed as in **a**. Supplementary Video 9 shows the complete video sequence. Insets show GFP fluorescence images at

re-opening. Mitotic cells expressing high levels of wild-type cyclin B fused to GFP underwent mitotic exit and cytokinesis upon flavopiridol treatment, but as expected, removing flavopiridol after 25 min never resulted in reversal of mitotic exit ($n = 6$) (Fig. 3c and Supplementary Video 9).

time 0. The level of wild-type cyclin B1 in **c** is approximately twice that of the non-degradable cyclin B1 expressed in **a** and **b**. **d**, Cells expressing high levels of non-degradable cyclin B1 are more likely to undergo mitotic exit reversal. Black diamonds indicate cells that did not separate chromatids upon treatment with flavopiridol, red diamonds show cells with separated chromatids. Open red diamonds indicate two cells that recondensed their chromosomes but had not opened their cleavage furrows by 90 min after flavopiridol removal. Error bars show s.e.m. **e**, A mitotic HeLa cell at metaphase (arrow) expressing non-degradable cyclin B1 has low levels of cyclin A expression, as assessed by immunofluorescence. Scale bars, 10 μm.

Cells expressing high levels of non-degradable cyclin B (as measured by GFP fluorescence) were more likely to undergo reversal than those expressing low levels (Fig. 3d). Previous work suggests that cell-cycle regulation in M phase shows the property of hysteresis, because entry into M phase from G2 requires higher Cdk1 activity

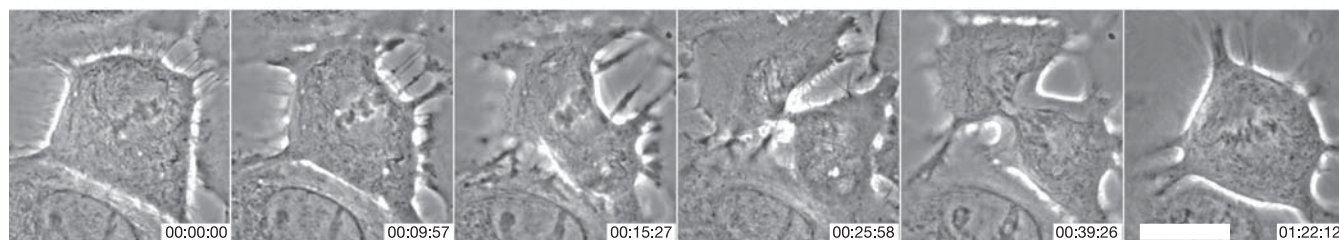


Figure 4 | Primary cultures of human cells can undergo mitotic exit reversal. An MG132-arrested metaphase human keratinocyte was imaged by phase-contrast microscopy after addition of 7.5 μM flavopiridol at time 0.

The flavopiridol was removed at 26 min. Supplementary Video 10 shows the complete video sequence. Scale bar, 10 μm.

than does maintaining M phase once it is achieved^{20,21}. This hysteresis might explain why many cells expressing relatively low levels of non-degradable cyclin B arrested at metaphase before flavopiridol treatment, but were unable to undergo reversal when flavopiridol was removed. Additionally, in confirmation of our previous conclusion that the reversal of mitotic exit and the re-establishment of high Cdk1 activity levels in the presence of proteasome inhibitor are probably driven by cyclin B and not cyclin A, we found that cells arrested at metaphase by expression of non-degradable cyclin B had nearly undetectable levels of cyclin A (Fig. 3e).

The above studies were all completed using immortalized cell lines. To determine whether these findings are broadly applicable to non-immortalized cells, we also tested primary human keratinocytes. We found that these cells, cultured in medium containing MG132, undergo mitotic exit and cytokinesis in response to flavopiridol treatment, and reverse back to metaphase when flavopiridol is removed (Fig. 4 and Supplementary Video 10).

The importance of regulated proteolysis for guiding the directionality of cell-cycle transitions is a central tenet of modern cell-cycle theory, extended from founding work on proteolysis of cyclins in early embryos^{1–4,22}. Here we have tested this idea and shown that if cyclin B is stabilized, the events of mitotic exit and cytokinesis are immediately reversible, even in cells that have progressed through chromosome decondensation, reconstitution of the interphase nucleus, mitotic spindle disassembly and cytokinesis to the stage of midbody formation. However, with increasing time after flavopiridol-induced mitotic exit, this reversibility wanes. Thus, cyclin degradation does provide directionality during mitotic exit, but this function must later be supplemented by other events that inhibit backtracking. These events are likely to involve other biochemical changes to Cdk1, including phosphorylation changes or the binding of small-molecule inhibitors. The ability to regulate mitotic exit in the absence of cyclin degradation in vertebrate cells will facilitate elucidation of the downstream regulators of the M to G1 phase transition in vertebrate cells.

METHODS

Cell culture and live cell imaging. *Xenopus* S3 cells stably expressing α -tubulin-GFP were grown at 23 °C in 70% L-15 medium supplemented with 15% fetal bovine serum. See Supplementary Video 12 for control images of these cells transiting M phase. HeLa cells were grown in DMEM medium with 10% FBS in 5% CO₂ at 37 °C. Normal human epidermal keratinocytes (Clonectics) were grown in defined keratinocyte SFM medium (Invitrogen). Time-lapse phase-contrast and fluorescence images were collected from cells grown on glass coverslips using a Zeiss Axiovert 200M microscope equipped with a Hamamatsu ORCA camera. When applicable, cells were incubated for 30–60 min with 25 μ M MG132 (Calbiochem) or with 100 ng ml⁻¹ nocodazole (Sigma). Flavopiridol was provided by the National Cancer Institute (Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis) and used at 5–10 μ M. In cells expressing cyclin B1-GFP, expression levels were quantified using Metamorph software (Molecular Devices).

Immunofluorescence and immunoblotting. Cells on coverslips were fixed in 2–3% formaldehyde in PHEM buffer (60 mM PIPES, 25 mM HEPES (pH 6.8), 10 mM EDTA, 4 mM MgCl₂) for 15 min, then permeabilized with 0.5% or 1% Triton X-100. HeLa cells in suspension were fixed in PHEM buffer containing 1.5% formaldehyde and 0.5% Triton X-100. Antibodies against *Xenopus* MKLP1 and aurora B were prepared in rabbits. Mouse antibody against human cyclin A was obtained from J. Gannon and T. Hunt, or obtained commercially (Abcam). Fluorescence images were taken using a Zeiss Axioplan II microscope with a Hamamatsu ORCA-II camera. Samples were treated with 4,6-diamidino-2-phenylindole (DAPI) and mounted in Vectashield containing 10 mM MgSO₄.

For immunoblotting, cells were lysed in NuPAGE western blotting sample buffer (Invitrogen) containing 50 mM dithiothreitol. Samples were electrophoresed on 4–12% gradient gels, transferred to PVDF membranes and probed with primary antibodies overnight at 4 °C. Blots were then treated with

appropriate HRP-conjugated secondary antibodies, and developed using SuperSignal West Pico chemiluminescent substrate (Pierce).

Received 30 December 2005; accepted 13 February 2006.

1. Minshull, J. *et al.* The role of cyclin synthesis, modification and destruction in the control of cell division. *J. Cell Sci.* **12** (suppl.), 77–97 (1989).
2. Reed, S. I. Ratchets and clocks: the cell cycle, ubiquitylation and protein turnover. *Nature Rev. Mol. Cell Biol.* **4**, 855–864 (2003).
3. Glotzer, M., Murray, A. W. & Kirschner, M. W. Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132–138 (1991).
4. Murray, A. W. & Kirschner, M. W. Cyclin synthesis drives the early embryonic cell cycle. *Nature* **339**, 275–280 (1989).
5. Clute, P. & Pines, J. Temporal and spatial control of cyclin B1 destruction in metaphase. *Nature Cell Biol.* **1**, 82–87 (1999).
6. Geley, S. *et al.* Anaphase-promoting complex/cyclosome-dependent proteolysis of human cyclin A starts at the beginning of mitosis and is not subject to the spindle assembly checkpoint. *J. Cell Biol.* **153**, 137–148 (2001).
7. den Elzen, N. & Pines, J. Cyclin A is destroyed in prometaphase and can delay chromosome alignment and anaphase. *J. Cell Biol.* **153**, 121–136 (2001).
8. Cross, F. R., Schroeder, L., Kruse, M. & Chen, K. C. Quantitative characterization of a mitotic cyclin threshold regulating exit from mitosis. *Mol. Biol. Cell* **16**, 2129–2138 (2005).
9. Parry, D. H. & O'Farrell, P. H. The schedule of destruction of three mitotic cyclins can dictate the timing of events during exit from mitosis. *Curr. Biol.* **11**, 671–683 (2001).
10. Sigrist, S., Jacobs, H., Stratmann, R. & Lehner, C. F. Exit from mitosis is regulated by *Drosophila* fizzy and the sequential destruction of cyclins A, B and B3. *EMBO J.* **14**, 4827–4838 (1995).
11. Visintin, R. *et al.* The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. *Mol. Cell* **2**, 709–718 (1998).
12. Chang, D. C., Xu, N. & Luo, K. Q. Degradation of cyclin B is required for the onset of anaphase in mammalian cells. *J. Biol. Chem.* **278**, 37865–37873 (2003).
13. Carmena, M. & Earnshaw, W. C. The cellular geography of aurora kinases. *Nature Rev. Mol. Cell Biol.* **4**, 842–854 (2003).
14. Glotzer, M. The molecular requirements for cytokinesis. *Science* **307**, 1735–1739 (2005).
15. Mishima, M., Pavicic, V., Gruneberg, U., Nigg, E. A. & Glotzer, M. Cell cycle regulation of central spindle assembly. *Nature* **430**, 908–913 (2004).
16. Losiewicz, M. D., Carlson, B. A., Kaur, G., Sausville, E. A. & Worland, P. J. Potent inhibition of CDC2 kinase activity by the flavonoid L86–8275. *Biochem. Biophys. Res. Commun.* **201**, 589–595 (1994).
17. Niiya, F., Xie, X., Lee, K. S., Inoue, H. & Miki, T. Inhibition of cyclin-dependent kinase 1 induces cytokinesis without chromosome segregation in an ECT2 and MgcRacGAP-dependent manner. *J. Biol. Chem.* **280**, 36502–36509 (2005).
18. Dranovsky, A. *et al.* Cdc2 phosphorylation of nucleolin demarcates mitotic stages and Alzheimer's disease pathology. *Neurobiol. Aging* **22**, 517–528 (2001).
19. Kramer, E. R., Scheuringer, N., Podtelejnikov, A. V., Mann, M. & Peters, J. M. Mitotic regulation of the APC activator proteins CDC20 and CDH1. *Mol. Biol. Cell* **11**, 1555–1569 (2000).
20. Pomerening, J. R., Sontag, E. D. & Ferrell, J. E. Jr. Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nature Cell Biol.* **5**, 346–351 (2003).
21. Sha, W. *et al.* Hysteresis drives cell-cycle transitions in *Xenopus laevis* egg extracts. *Proc. Natl Acad. Sci. USA* **100**, 975–980 (2003).
22. Evans, T., Rosenthal, E. T., Youngblom, J., Distel, D. & Hunt, T. Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell* **33**, 389–396 (1983).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature. A summary figure is also included.

Acknowledgements We thank the staff at the OMRF Imaging Center and W. Martin for technical assistance. We thank Sanofi-Aventis Pharmaceuticals and the National Cancer Institute for providing flavopiridol. We thank J. Pines for providing cyclin B plasmids and P. Davies for TG-3 anti-phospho-nucleolin antibody. We thank J. Gannon and T. Hunt for antibodies against Cdk1 and cyclins. This work was supported by grants from the National Institutes of Health (to G.J.G. and P.T.S.), from the American Cancer Society (to P.T.S.) and from an NIH training grant (to D.L.S.).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to G.J.G. (GJG@omrf.ouhsc.edu).

Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR

Adam Lange¹, Karin Giller¹, Sönke Hornig², Marie-France Martin-Eauclaire³, Olaf Pongs², Stefan Becker¹ & Marc Baldus¹

The active site of potassium (K^+) channels catalyses the transport of K^+ ions across the plasma membrane¹—similar to the catalytic function of the active site of an enzyme—and is inhibited by toxins from scorpion venom. On the basis of the conserved structures of K^+ pore regions² and scorpion toxins^{3,4}, detailed structures for the K^+ channel–scorpion toxin binding interface have been proposed. In these models and in previous solution-state nuclear magnetic resonance (NMR) studies using detergent-solubilized membrane proteins^{5,6}, scorpion toxins were docked to the extracellular entrance of the K^+ channel pore assuming rigid, preformed binding sites^{7–13}. Using high-resolution solid-state NMR spectroscopy, here we show that high-affinity binding of the scorpion toxin kaliotoxin to a chimaeric K^+ channel (KcsA-Kv1.3)^{14,15} is associated with significant structural rearrangements in both molecules. Our approach involves a combined analysis of chemical shifts and proton–proton distances and demonstrates that solid-state NMR is a sensitive method for analysing the structure of a membrane protein–inhibitor complex. We propose that structural flexibility of the K^+ channel and the toxin represents an important determinant for the high specificity of toxin– K^+ channel interactions.

Solid-state NMR (ssNMR) represents a spectroscopic method to study molecular structure and dynamics in a non-crystalline environment and has been used to probe individual constraints in tailored isotope-labelled membrane proteins (see, for example, refs 16, 17). Here, we show how two-dimensional ssNMR spectroscopy can be used to study an entire membrane protein complex and its uniformly labelled constituents. First, we recorded ssNMR spectra for proteoliposomes containing purified KcsA-Kv1.3 in complex with uniformly labelled [^{13}C , ^{15}N]kaliotoxin (KTX). *De novo* sequential resonance assignments for 32 out of 38 KTX residues were obtained using a series of two-dimensional (^{13}C , ^{15}N) and (^{13}C , ^{13}C) correlation experiments¹⁸ (see Supplementary Information). As indicated in a two-dimensional [^{13}C , ^{13}C] correlation spectrum (shown in green in Fig. 1a–c), the ^{13}C line widths seen for isotope-labelled KTX ranged between 0.7–0.8 p.p.m., suggesting that the KcsA-Kv1.3 toxin receptor site is homogeneous and well defined. A comparison to two-dimensional [^{13}C , ^{13}C] correlations observed for unbound, solid-phase KTX¹⁸ (shown in red in Fig. 1a–c; see also Supplementary Information) revealed significant ^{13}C chemical shift changes for several residues. A secondary chemical shift analysis¹⁹ of our ssNMR data showed that bound and free KTX had a comparable backbone fold (see Supplementary Information). We used this information for assignments of CHHC²⁰ correlations that are sensitive to proton–proton distances and hence to the three-dimensional structure of KTX. Our analysis yielded a structural model of KTX with a backbone root mean square deviation (r.m.s.d.) of 1.1 Å that

deviated from the structure of free KTX by a backbone r.m.s.d. of 2.6 Å (Fig. 2; see also Supplementary Information).

Previous mutagenesis studies on KTX^{9,10} and agitoxin 2 (ref. 21) identified three toxin– K^+ channel interface regions: at one side of the α -helix (S11 and L15), along the second β -strand including the following turn (R24 to R31) and at the end of the third β -strand (T36 and P37) (Fig. 2a). Mutation of residues R24 and K27—critical for KTX binding to Kv1.3 channels^{9,10}—also significantly decreased KTX binding to KcsA-Kv1.3 (90-fold and 180-fold, respectively), indicating a similar functional interaction surface (Fig. 3). We

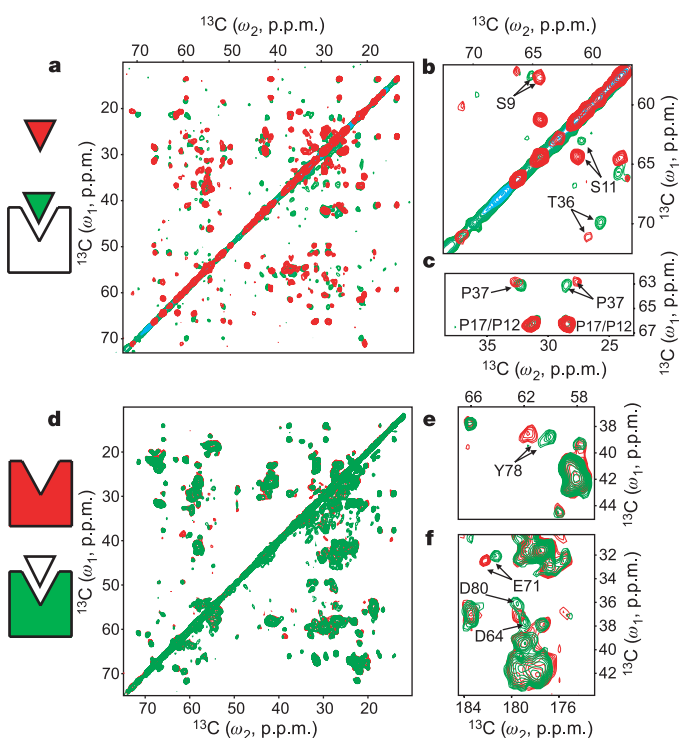


Figure 1 | Comparison of ssNMR data of KTX and KcsA-Kv1.3 in free and complex form. **a–c**, (^{13}C , ^{13}C) Two-dimensional spin diffusion spectrum obtained on uniformly (U) [^{13}C , ^{15}N]-labelled KTX in free, solid-phase form (red) and in complex with KcsA-Kv1.3 (green). KTX is symbolized by triangles at the left of the panel. **d–f**, Two-dimensional spin diffusion spectrum obtained on U- ^{13}C , ^{15}N]KcsA-Kv1.3 (red) and KTX-U- ^{13}C , ^{15}N]KcsA-Kv1.3 (green, see cartoon at the left of the panel). In both data sets, spectral regions are shown for specific residue types (**b**, **c** and **e**, **f**). Carbon resonance frequencies ω_1 and ω_2 are in p.p.m.

¹Max Planck Institute for Biophysical Chemistry, Department of NMR-Based Structural Biology, 37077 Göttingen, Germany. ²Universität Hamburg, Zentrum für Molekulare Neurobiologie, Institut für Neuronale Signalverarbeitung, Falkenried 94, 20251 Hamburg, Germany. ³CNRS, Institut Jean Roche, Université de la Méditerranée, Bd Pierre Dramard, 13916 Marseille Cedex 20, France.

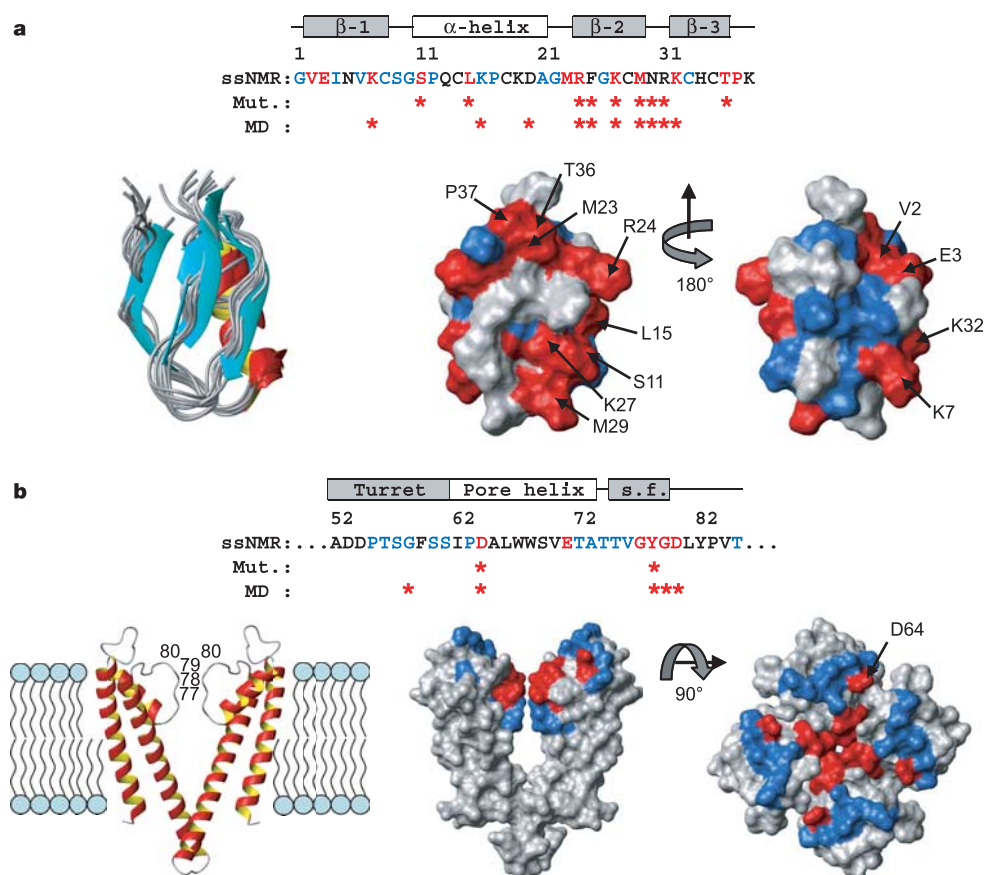


Figure 2 | Analysis of ssNMR data for KTX-KcsA-Kv1.3. **a, b,** Amino acid sequence of KTX (**a**) and of the KcsA-Kv1.3 pore region (**b**) with secondary structures (boxes, s.f. stands for selectivity filter). Residues characterized by significant chemical-shift changes upon complex formation are coloured red; those involved according to mutagenesis (Mut.)^{9,10} and molecular dynamics (MD)¹³ are marked with an asterisk. Blue, unperturbed residues;

black, not determined. Perturbed (red) or unperturbed (blue) residues are indicated in surface plots of KTX (**a**, ssNMR structure of KTX bound to KcsA-Kv1.3) and channel (**b**, Protein Data Bank 1K4C). In addition to the KTX structures bound to KcsA-Kv1.3 (**a**), models of two and four membrane-inserted KcsA subunits are included (**b**).

observed in the three interface regions significant ssNMR chemical-shift changes for KTX residues S11 and L15 (α -helix), M23, R24, K27 and M29 (second β -strand), and T36 and P37 (end of third β -strand) (Fig. 2a). In a surface plot (Fig. 2a), these residues are found on one side of the KTX three-dimensional structure in complex with the channel, consistent with the view that these KTX residues define a direct interaction surface with the KcsA-Kv1.3 channel. On the opposite side of the interaction surface, residues V2, E3 and K7 also exhibited significant chemical-shift changes, reflecting indirect effects possibly due to conformational changes involving β -sheet contacts between the first and third β -strand⁹. Mutations of V2, E3 and K7 mildly affected both toxin inhibition of Kv1.3 currents (half-maximal inhibition concentration, IC₅₀) expressed in *Xenopus* oocytes and toxin affinity to KcsA-Kv1.3 protein (dissociation constant, K_d) in competition binding assays (Fig. 3a, b). These residues—as for the interaction of charybdotoxin with the Shaker channel²²—are not critical but are apparently influential in the toxin-K⁺ channel interaction.

Next, we recorded ssNMR spectra for the isotope-labelled KcsA-Kv1.3 channel at high (50 mM) and low (≤ 1 mM) KCl concentration. Analysis of these two-dimensional ssNMR data and the observed ¹³C line widths indicated a well-folded, largely α -helical membrane protein, in agreement with the crystal structure of conductive KcsA²³ (see below). Solid-state NMR measurements in the presence (green) or absence (red) of unlabelled KTX (Fig. 1d–f; see also Supplementary Information) demonstrated that the structure of outer and inner pore helices is largely maintained in KcsA-Kv1.3 after

binding KTX. By contrast, chemical-shift changes were seen for channel residues in the KTX-binding region. Its non- α -helical structure facilitated ssNMR resonance assignments (see Fig. 2b). We observed significant chemical-shift changes for residues in both the pore helix and the selectivity filter (I62–D80, indicated in red; Fig. 2b). The ssNMR data directly show that D64 in the vestibule of KcsA-Kv1.3 represents an important interaction site for KTX (Fig. 3). Mutation of the homologous residue in Kv1.3 channels reduces KTX affinity by more than three orders of magnitude^{9,10}. Large (>1.5 p.p.m.) chemical-shift changes were seen for G77 NH, Y78 CA and G79 NH in the selectivity filter. Further evidence for a conformational change in the filter region came from an NHHc²⁰ experiment that indirectly encodes proton–proton proximities. Also, we observed chemical-shift changes for the side chains of E71 and D80 that form carboxyl–carboxylate pairs on the back side of the filter²³.

In crystal structures^{23–25}, the KcsA selectivity filter adopts a conductive conformation at high K⁺ concentration. The active site of the K⁺ channel can take up different conformations in response to ion binding²⁶. When Na⁺ replaces K⁺ in solution, the selectivity filter undergoes a conformational change. As a result, the filter collapses shut²⁶. This is reflected by large changes in backbone angles of filter residues, indicating a remarkable conformational flexibility in the active site of the ion channel^{23–25}. Using the distinct relationship between protein backbone structure and ssNMR chemical shift, we determined the backbone torsion angles for filter residues T75–G79 (summarized in Fig. 4). The results show that the backbone conformation of the KcsA-Kv1.3 selectivity filter in the unbound form

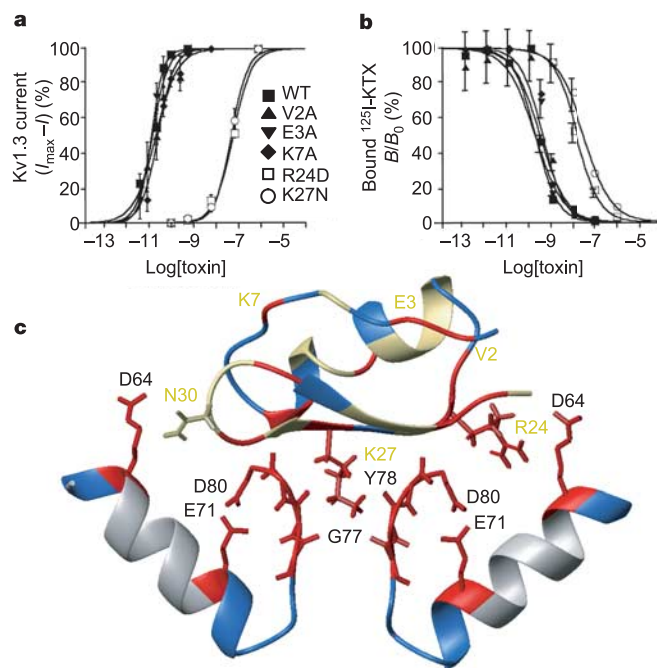


Figure 3 | Competition binding assays of KTX mutants and structural model of the KTX-KcsA-Kv1.3 complex. **a**, Inhibition of Kv1.3-mediated outward currents by wild-type (WT) KTX and KTX mutants. **b**, Binding of ¹²⁵I-KTX to purified KcsA-Kv1.3 protein in the absence (*B*₀) or presence (*B*) of toxin (wild-type KTX and KTX mutants). Symbols are as in **a**. Error bars indicate s.e.m. **c**, Model of the KTX-KcsA-Kv1.3 complex. Residues (yellow labels, toxin; black labels, channel) affected or unperturbed by complex formation (according to solid-state NMR chemical-shift mapping) are indicated in red and blue, respectively (see also Fig. 2).

agrees with the conductive conformation of KcsA. For the KTX-bound channel we obtained torsion angles that resemble those of the conductive conformation of KcsA for filter residues 75–77; however, for filter residues 78 and 79 we obtained torsion angles that approximate the collapsed structure of KcsA. The data indicate that in the presence of KTX, the KcsA-Kv1.3 selectivity filter adopts a novel conformation with features of both the conducting and the collapsed conformation, as seen in KcsA crystal structures²³ (Fig. 4). Obviously, the ability of the selectivity filter to undergo particular conformational changes enables KcsA-Kv1.3 to bind KTX with high affinity.

Through the combination of structural and functional data we derived a model for the KTX-KcsA-Kv1.3 complex (Fig. 3c). The lysine (K27) side chain, which is highly conserved in scorpion toxins, protrudes from the interaction surface and plugs the K⁺ channel pore^{3,4}. The positively charged K27 ε-amino group comfortably fits into the selectivity filter near G77, and the hydrophobic methylene groups of the inserted K27 side chain replace water molecules in the entry region of the pore. Thus, the K27 side chain makes intimate contact to a pore region, the backbone of which is highly K⁺ sensitive and has a different conformation in the conductive versus non-conductive state²³. This view corroborates observations that K27 toxin variants become sensitive to the K⁺ channel's gating conformation²⁷ and that mutations of the Y78 side chain strongly affect the K⁺ sensitivity of toxin binding to the Shaker channel, although our data show that the Y78 side chain is not directly involved in toxin binding²¹. Furthermore, the model also explains observations that toxin association rate is sensitive to the K⁺ channel's gating conformation, but dissociation rate is not²⁸.

Mutagenesis studies^{9,29} have indicated that residues in the turret of the P-loop region may contribute to the toxin-K⁺ channel interaction surface; for example, Kv1.3 residue G380 (refs 9, 10)

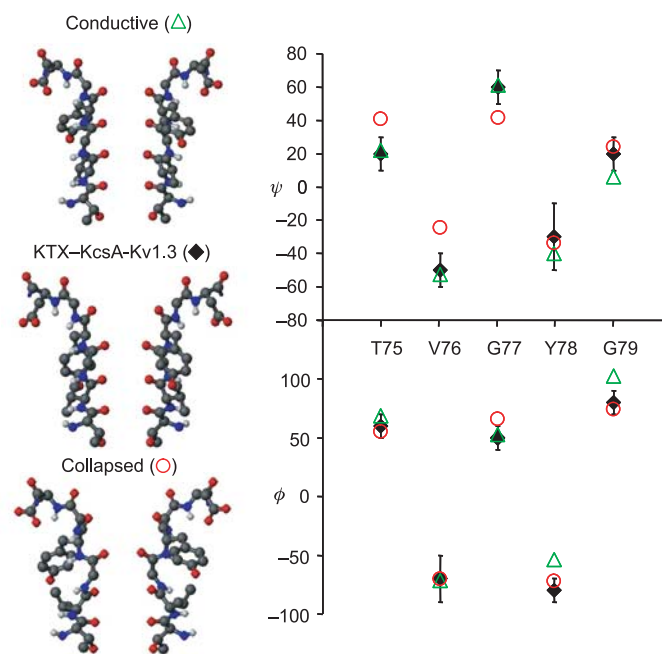


Figure 4 | Comparison of the selectivity filter in the KTX-KcsA-Kv1.3 complex derived from ssNMR data to KcsA X-ray structures. Dihedral angles Ψ and Φ for selectivity-filter residues T75–G79 derived from ssNMR chemical shifts (black diamonds, error bars represent systematic errors) of KcsA-Kv1.3 in complex with KTX. Values are compared to those in KcsA X-ray structures²³ obtained at high (green triangles) and low (red circles) K⁺ concentrations. Data were used to build the selectivity-filter structure shown for two subunits (left panel) together with the conductive (Protein Data Bank 1K4C) and collapsed (Protein Data Bank 1K4D) structure of KcsA.

(equivalent to G58 in Fig. 2b). Alternatively, bulky substitutions at this glycine residue²² in the turret region may simply hinder free access of toxin to the channel binding site without directly influencing the toxin-K⁺ channel interaction. Indeed, we did not observe chemical-shift variations indicative of side-chain interaction(s) of the KTX-KcsA-Kv1.3 complex in this channel region. These findings support our structural model that high-affinity toxin binding is characterized by a deeper insertion of the toxin into the selectivity filter than previously thought. Our data also show that the insertion is associated with conformational changes in both molecules. It is likely that this is a prerequisite for high-affinity binding. Then, the intrinsic dynamics of the selectivity filter^{23–25} provide the basis for formation of a tight complex with the active site of the K⁺ channel.

Our ssNMR results show that the interaction between a scorpion toxin and a K⁺ channel is associated with structural rearrangements in different segments of the ligand and the ligand-binding site, including the selectivity filter of the channel. Accounting for these conformational changes may be crucial for a full thermodynamic characterization of high-affinity ligand binding to an ion channel. The application of ssNMR spectroscopy presented here can be exploited to advance current structure-based design studies, leading to novel therapeutic agents in K⁺ pharmacology.

METHODS

Preparation of KTX and the KTX-KcsA-Kv1.3 complex. [¹³C, ¹⁵N]-labelled KTX was produced as an intein fusion protein. Expression and purification of the KcsA-Kv1.3 chimaeric channel was based on previous work¹⁴ using an optimized pQE32 expression construct with the additional mutation L90C (see Supplementary Information).

Electrophysiological experiments and binding studies. Kv1.3 channels were expressed in *Xenopus* oocytes. Kv1.3-mediated outward currents were measured at +60 mV test potentials. Kaliotoxin and kaliotoxin mutants were applied by bath perfusion to determine IC₅₀ from current amplitudes measured in the

absence ($I_{\max}$) and presence of toxin (I). Current inhibition was calculated as $I_{\max} - I$ in per cent ($n = 3-6$). Competition of ^{125}I -KTX binding to purified KcsA-Kv1.3 protein by addition of increasing concentrations of wild-type KTX or the mutants KTX(V2A), KTX(E3A), KTX(K7A), KTX(R24D) or KTX(K27N) was assayed as described¹⁴.

Solid-state NMR experiments and resonance assignments. Two-dimensional ssNMR experiments were conducted on a 14.1 T (^1H resonance frequency: 600 MHz) wide-bore instrument (Bruker Biospin) equipped with a 4-mm triple-resonance (^1H , ^{13}C , ^{15}N) MAS probe. All experiments were carried out at probe temperatures of -25 to -30°C . Pelleted proteoliposomes were transferred to a 4-mm MAS rotor. The MAS frequency was set to 9.375 kHz. For the sequential assignment of $^{13}\text{C}/^{15}\text{N}$ backbone and side-chain resonances, the homonuclear and heteronuclear experiments described in the Supplementary Information were performed. ^{13}C and ^{15}N resonances were calibrated using adamantane and the tripeptide AGG as an external reference. In Figs 2 and 3, residues that exhibit chemical shift changes >0.7 p.p.m. (that is, larger than the observed line width) for at least one of the resonances $\text{C}\alpha$, $\text{C}\beta$, $\text{C}\gamma$, $\text{C}\delta$, or $\text{C}\epsilon$ are indicated in red. Residues shown in black in Fig. 2 are characterized by incomplete ssNMR resonance assignments. For KcsA-Kv1.3, NH chemical shift changes were also considered. For the determination of the tertiary fold of bound KTX, a CHHC²⁰ experiment was performed. In addition to the analysis of conformation-dependent chemical shifts, a two-dimensional NHHC²⁰ spectrum was used to infer the conformation of the selectivity filter of KcsA-Kv1.3 (see Supplementary Information).

Structure calculations. Backbone chemical shifts were analysed in terms of backbone protein structure using published software routines. Proton-proton distance constraints were obtained by comparison of experimental and simulated CHHC spectra using the three-dimensional structure of unbound KTX as a reference.

Received 8 September 2005; accepted 15 February 2006.

- Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, Massachusetts, 2001).
- Long, S. B., Campbell, E. B. & MacKinnon, R. Crystal structure of a mammalian voltage-dependent Shaker family K^+ channel. *Science* **309**, 897–903 (2005).
- Miller, C. The charybdotoxin family of K^+ channel-blocking peptides. *Neuron* **15**, 5–10 (1995).
- Garcia, M. L., Gao, Y. D., McManus, O. B. & Kaczorowski, G. J. Potassium channels: from scorpion venoms to high-resolution structure. *Toxicon* **39**, 739–748 (2001).
- Takeuchi, K. et al. Structural basis of the KcsA K^+ channel and agitoxin2 pore-blocking toxin interaction by using the transferred cross-saturation method. *Structure* **11**, 1381–1392 (2003).
- Yu, L. et al. Nuclear magnetic resonance structural studies of a potassium channel-charybdotoxin complex. *Biochemistry* **44**, 15834–15841 (2005).
- Mackinnon, R. & Miller, C. Mutant potassium channels with altered binding of charybdotoxin, a pore-blocking peptide inhibitor. *Science* **245**, 1382–1385 (1989).
- Hidalgo, P. & Mackinnon, R. Revealing the architecture of a K^+ channel pore through mutant cycles with a peptide inhibitor. *Science* **268**, 307–310 (1995).
- Aiyar, J. et al. Topology of the pore-region of a K^+ channel revealed by the NMR-derived structures of scorpion toxins. *Neuron* **15**, 1169–1181 (1995).
- Aiyar, J., Rizzi, J. P., Gutman, G. A. & Chandy, K. G. The signature sequence of voltage-gated potassium channels projects into the external vestibule. *J. Biol. Chem.* **271**, 31013–31016 (1996).
- MacKinnon, R., Cohen, S. L., Kuo, A. L., Lee, A. & Chait, B. T. Structural conservation in prokaryotic and eukaryotic potassium channels. *Science* **280**, 106–109 (1998).
- Eriksson, M. A. L. & Roux, B. Modeling the structure of Agitoxin in complex with the Shaker K^+ channel: A computational approach based on experimental distance restraints extracted from thermodynamic mutant cycles. *Biophys. J.* **83**, 2595–2609 (2002).
- Yu, K. et al. Computational simulations of interactions of scorpion toxins with the voltage-gated potassium ion channel. *Biophys. J.* **86**, 3542–3555 (2004).
- Legros, C. et al. Generating a high affinity scorpion toxin receptor in KcsA-Kv1.3 chimeric potassium channels. *J. Biol. Chem.* **275**, 16918–16924 (2000).
- Legros, C. et al. Engineering-specific pharmacological binding sites for peptidyl inhibitors of potassium channels into KcsA. *Biochemistry* **41**, 15369–15375 (2002).
- Creuzet, F. et al. Determination of membrane-protein structure by rotational resonance NMR: Bacteriorhodopsin. *Science* **251**, 783–786 (1991).
- Opella, S. J. & Marassi, F. M. Structure determination of membrane proteins by NMR spectroscopy. *Chem. Rev.* **104**, 3587–3606 (2004).
- Lange, A. et al. A concept for rapid protein-structure determination by solid-state NMR spectroscopy. *Angew. Chem. Int. Edn Engl.* **44**, 2089–2092 (2005).
- Luca, S. et al. Secondary chemical shifts in immobilized peptides and proteins: A qualitative basis for structure refinement under Magic Angle Spinning. *J. Biomol. NMR* **20**, 325–331 (2001).
- Lange, A., Luca, S. & Baldus, M. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.* **124**, 9704–9705 (2002).
- Ranganathan, R., Lewis, J. H. & MacKinnon, R. Spatial localization of the K^+ channel selectivity filter by mutant cycle-based structure analysis. *Neuron* **16**, 131–139 (1996).
- Goldstein, S. A. N., Pheasant, D. J. & Miller, C. The charybdotoxin receptor of a Shaker K^+ channel: Peptide and channel residues mediating molecular recognition. *Neuron* **12**, 1377–1388 (1994).
- Zhou, Y., Morais-Cabral, J. H., Kaufman, A. & MacKinnon, R. Chemistry of ion coordination and hydration revealed by a K^+ channel-Fab complex at 2.0 Å resolution. *Nature* **414**, 43–48 (2001).
- Zhou, Y. & MacKinnon, R. The occupancy of ions in the K^+ selectivity filter: Charge balance and coupling of ion binding to a protein conformational change underlie high conduction rates. *J. Mol. Biol.* **333**, 965–975 (2003).
- Lenaeus, M. J., Vamvouka, M., Focia, P. J. & Gross, A. Structural basis of TEA blockade in a model potassium channel. *Nature Struct. Mol. Biol.* **12**, 454–459 (2005).
- Gouaux, E. & MacKinnon, R. Principles of selective ion transport in channels and pumps. *Science* **310**, 1461–1465 (2005).
- Park, C. S. & Miller, C. Interaction of charybdotoxin with permeant ions inside the pore of a K^+ channel. *Neuron* **9**, 307–313 (1992).
- Anderson, C. S., Mackinnon, R., Smith, C. & Miller, C. Charybdotoxin block of single Ca^{2+} -activated K^+ channels - Effects of channel gating, voltage, and ionic-strength. *J. Gen. Physiol.* **91**, 317–333 (1988).
- Gross, A., Abramson, T. & Mackinnon, R. Transfer of the scorpion toxin receptor to an insensitive potassium channel. *Neuron* **13**, 961–966 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to H. Kratzin and K. Overkamp for peptide analysis. Technical assistance by B. Angerstein is gratefully acknowledged. We thank C. Griesinger for continuous support of this project. This work was funded in part by the DFG and by a PhD. fellowship to A.L. from the Stiftung Stipendien-Fonds of the Verband der Chemischen Industrie.

Author Contributions O.P., S.B. and M.B. are co-senior authors.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.B. (maba@mpibpc.mpg.de).

-  FOCUS
-  SPOTLIGHT
-  RECRUITMENT
-  ANNOUNCEMENTS
-  EVENTS

naturejobs

What now for biotechnology?

To celebrate 30 years of the biotechnology industry, analysts Ernst & Young asked five leaders from academia, industry and government to forecast the sector's future. Three of the five predict that solving health and nutrition problems in the developing world will create an opportunity for investment and jobs

To feed the world's growing population, agriculture over the next 20 years will have to double its production from the same amount of land, says Nobel laureate Norman Borlaug, the 'father' of the green revolution. Two approaches can help, he says: continue to work on 'producer-oriented' traits, such as tolerance to temperatures and drought, and do more work to genetically modify plants to improve their nutritional and health benefits.

As well as nutrition problems, developing countries must deal with infectious diseases such as tuberculosis and malaria. Public-private partnerships are now pouring money into this problem, says Victoria Hale, chief executive of the Institute for OneWorld Health, a non-profit pharmaceutical company in San Francisco, California. But more needs to be done to make biotech products that can control these

diseases available to the poor. This is a great challenge for young scientists — assuming that the funding materializes.

Finally, as countries develop, they may fall prey to the downsides of Western progress: "the diseases of a richer society", such as cancer, heart disease and diabetes, says David Weatherall of the University of Oxford. Drugs already exist in the West to treat these diseases, but costs need to come down to make them accessible to the developing world. And better, cheaper diagnostics will help physicians tackle the conditions earlier and with better targeted drugs in all parts of the world, says Weatherall.

These three challenges could lead to more funding and jobs, but that depends on the developed countries choosing to address the problems — and convincing them to do just that is perhaps the biggest challenge of all.



Paul Smaglik, Naturejobs editor

CONTACTS

Publisher: Ben Crowe
Editor: Paul Smaglik
Assistant Editor: Corie Lok

European Head Office, London
 The Macmillan Building,
 4 Crinan Street
 London N1 9XW, UK
 Tel: +44 (0) 20 7843 4961
 Fax: +44 (0) 20 7843 4996
 e-mail: naturejobs@nature.com

European Sales Manager:
 Andy Douglas (4975)
Natureevents: Silke Opstrup (4994)

UK/RoW/Ireland/Italy:

Nils Moeller (4953)
 Irene Viglia-Atton (4944)
Scandinavia/Spain/Portugal:
 Evelina Rubio Håkansson (4973)
France/Switzerland/Belgium:
 Amelie Pequignot (4974)
**Germany/Austria/
 The Netherlands:**
 Reya Silao (4970)

Advertising Production Manager:

Billie Franklin
 To send materials use London
 address above.
 Tel: +44 (0) 20 7843 4814

Fax: +44 (0) 20 7843 4996
 e-mail: naturejobs@nature.com
Naturejobs web development: Tom Hancock
Naturejobs online production: Niamh Shields

**European Satellite Office
 Germany/Austria/
 The Netherlands:**
 Patrick Phelan
 Tel: +49 89 54 90 57 11
 Fax: +49 89 54 90 57 20
 e-mail: p.phelan@nature.com

US Head Office, New York
 75 Varick Street,
 9th Floor, New York,

NY 10013-1917
 Tel: +1 800 989 7718
 Fax: +1 800 989 7103
 e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo
 Chiyoda Building,
 2-37 Ichigayatamachi,
 Shinjuku-ku,
 Tokyo 162-0843
 Tel: +81 3 3267 8751
 Fax: +81 3 3267 8746
Asia-Pacific Sales Director: Rinoko Asami
 e-mail: rasami@naturejpn.com

Mentoring mismatch



Is your adviser not the role model or mentor of your dreams? Then take charge of the situation and find the right people. **Kendall Powell** plays matchmaker.



Paul's* graduate supervisor stood beside his bench yelling at him about a failed experiment. Sadly, this was nothing new. But Paul knew that they had reached breaking point when he chased his supervisor into the hallway to continue the shouting match.

"Once you go that far, there's no mentoring going on and no good way back," says Paul, now an assistant professor at a California research institute. Paul took an extreme option — the only sensible one in his case — and switched labs in the middle of his graduate career. He found a mentor who was a valuable source of guidance as Paul made the transition from postgraduate to investigator.

Fortunately, such drastic measures are not the most common solution for bad mentoring matches. A milligram of prevention is worth a kilogram of cure when it comes to communicating your goals and expectations (see *Nature* **422**, 784–785; 2003).

Most trainees need several different mentors to serve as role models and to advise them on scientific and career questions. To help meet this need, a growing number of universities are introducing policies that match trainees with mentors in addition to their advisers.

For example, several UK universities have put 'backstop' policies in place for doctoral students in which a secondary adviser acts as a sounding board for the student's concerns. As there is no possible conflict of interest in their publication record or advancement,

Keith Micoli: if you feel uncomfortable bringing up career aspirations during an interview, it won't get easier later.

a secondary adviser or an external mentor can play a key role.

Although advisers can't be all things to their trainees, yours may well give you career advice — but you'll need to ask, says Sunita Jones, a former postdoc and peer counsellor at Stanford University in Palo Alto, California. She says postdocs should not feel like they are under their adviser's control.

"You already have your degree, no one can take it away. This training is for your future and you are in control," says Jones, now a research administrator at St Mary's Hospital in Manchester, UK. If something is not working for you, tell your adviser, who, she says, should also be a colleague by this point.

Aim for comfort

Equally important is starting off on the right foot with a mentor at the earliest stages, says Keith Micoli, chairman of the National Postdoctoral Association (NPA) board and research associate at the University of Alabama, Birmingham.

"You should talk to an investigator before you even accept a lab position," says Micoli. And you should make your career aspirations clear, he adds. "If you don't feel comfortable during the interview to bring it up, it's not going to get more comfortable later." You can also research the NPA's postdoc policy database to find the postdoctoral infrastructure available at more than 100 institutions.

Once you have accepted a position, set out scientific and career goals and reasonable mentoring expectations with your supervisor as soon as possible. Doing so will show you the areas your adviser can

handle and where you'll need a backup. A good place to start is with an individual development plan (IDP), a written framework for regular discussions with an adviser on your scientific and career development.

"If you don't map a course, then you don't know whether you are making progress or running in circles," says Micoli. He notes that IDPs can help introduce difficult topics, such as how much of a research project a postdoc can expect to take with them to their next academic position. IDP-type programmes are being implemented by universities including Stanford and Vanderbilt to improve postdoctoral mentoring. Vanderbilt School of Medicine in Nashville, Tennessee, has proposed that postdocs write an IDP upon arrival, update it at six months, and then redraft it with an adviser each year at re-appointment. Students can also benefit from these regular career-development discussions.

Alexandra could have used just such a system when her graduate supervisor, Tom, went missing for weeks at a time to deal with personal problems. "I was literally doing everything myself," she says. "But when it came to writing up my thesis, I felt totally unprepared."

Luckily, she found a mentor in a former lab supervisor. "He really saved me," she says, of the man who helped her revise her thesis and set a defence date. Yet she notes that Tom was a good scientific mentor, pushing her to do rigorous controls and repeat experiments. Her former lab adviser, on the other hand, was not a good match on scientific aspects.

Hunting down help

Like Alexandra, trainees should recognize their adviser's mentoring strengths and weaknesses, and seek out others who can fill the gaps. International students would do well to find someone to teach them the ropes of a new science system, says Rosana Kapeller, who left Brazil for the United States to do a PhD. Young investigators may not have the time or skills to do this, so look to your peers or someone more established.

Kapeller, now vice-president for research at Renegade Therapeutics in Cambridge, Massachusetts, found help from her department chair. "He helped me navigate the difficult obstacles as a young scientist," she says. Students need to recognize that they may not get everything from their own adviser, says Kapeller.

Jones suggests finding someone more neutral than an adviser as you ponder different careers. "Go across the hall and find a mentor who's going to listen — somebody to bounce things off, someone who won't judge," she says. During her own job search, she created a postdoc transition group to relieve the pressure and stress. The group met every month over lunch to share stories and contacts. When members landed positions — as they all eventually did — they told each other what had worked for them.

Beyond department



Sunita Jones: find someone outside to bounce ideas off.

colleagues, you may want to find someone outside your institution for discussing sensitive subjects or to fill a specific mentoring role. MentorNet is a service that connects college, graduate, postdoc and new faculty protégés with mentors through a one-to-one e-mail system. Last year, about 700 graduate students and 70 postdocs connected with mentors through the non-profit programme.

"An adviser's advice is invaluable, but not always exhaustive," notes Carol Muller, founder and chief executive of MentorNet in San Jose, California. "These are people who can give perspectives and information about alternative career paths or the experiences of women and minorities in science." They can address future plans, doubts and fears you might not want known at your home institution — the timing of having children, dealing with dual-career families, when to publish and move on.

"This unbiased individual is not grading you or sitting on a dissertation or tenure committee. It is completely appropriate to reach out and seek a variety of sources of advice at this stage — these people are the beginnings of a professional network," says Muller.

Perfect match

MentorNet can also match you up with a role model — as in the case of one young engineer who had never met another female, African-American, engineering professor. The programme matches people on the basis of ranked preferences such as gender, ethnicity, institution, location and discipline.

Networkers could also contact interesting seminar speakers by e-mail and join groups such as Women Entrepreneurs in Science and Technology (WEST). Kapeller, a WEST board member, says the best way to make lasting connections is to volunteer for committees with regular small-group meetings.

For exploring alternative career paths that your adviser may not be familiar with, call human-resources departments and ask what they are looking for or search out friends of friends already working in a sector for an informational interview. Alexandra took a 'self-help' route, reading books on handling work stress and communication problems, talking to an occupational therapist, and hiring a career coach.

Jones, Kapeller and others say that good mentoring relationships last a lifetime, get you through big life transitions and prepare you to become a mentor for others. Kapeller says she enjoys shedding light to help people find their own path. "It's about forging relationships and that's always a two-way street."

Even with excellent outside advice, remember that your adviser will continue to be an adviser — sometimes an even better one — when you leave the lab. Paul's second graduate adviser became a valuable confidant once he became a postdoc. And it's likely you will find yourself in your adviser's shoes someday.

"If you look at things from the investigator's perspective and see how you expect them to behave, it's completely different," says Paul. "I never thought I'd be one of those investigators who asks for data when travelling, but the other day I told my technician to send me a text message as soon as my plane landed!"

Kendall Powell is a freelance science writer in Broomfield, Colorado.

*Some names have been changed to protect privacy.



Carol Muller: a network can grow from several advisers.

WEB LINKS

FASEB's IDP

♦ opa.faseb.org/pdf/idp.pdf

MentorNet

♦ www.mentornet.net

NPA's postdoctoral policy database

♦ www.nationalpostdoc.org/policy/institutional

MOVERS

Nikolaus Rajewsky, head of bioinformatics research, Max Delbrück Center, Berlin, Germany



2003-06: Assistant professor, Center for Comparative Functional Genomics, New York University.

2002-03: Research assistant professor, Rockefeller University, New York.

1999-2002: Postdoc, Rockefeller University, New York.

1998-99: Postdoc, Rutgers University, New Jersey.

"What field of science you belong to depends on what questions you're asking," says Nikolaus Rajewsky, who is formally trained as a physicist. "Having said this, I am a biologist." Rajewsky, who joins the Max Delbrück Center (MDC) in Berlin next month, has frequently reformulated his questions. This is why his career has had such a broad span: mathematics, informatics and biology.

His questions at his current base of New York University have focused on 'microRNAs' — short pieces of RNA that help regulate vital pathways such as insulin secretion and cholesterol synthesis, but whose mechanisms of action are still not understood. That is as hard-core as biology gets — but biology was the first subject Rajewsky dropped in school.

After a PhD in theoretical physics from the University of Cologne (where he also took a master's in piano-playing), Rajewsky took a summer course at Princeton University called 'biology for mathematicians' in 1999. "It was then that I decided on biology," he says. He took up a postdoc position investigating regulatory elements in different genomes at Rockefeller University. "At that time, I knew little more about biology than the double helix and Darwin's theories," he says.

Now a much-courted expert in bioinformatics, Rajewsky taught himself biology — from books, seminars and the "inspiring" scientific environment of Rockefeller.

"The questions I'm asking are biological ones," he says, "and the tools I'm using are mostly informatics and wet-lab experiments." He advises every biology student to get a solid education in bioinformatics: "You'll need it."

His new appointment also covers Berlin's Charité medical school. "I like the possibility of bringing basic science to clinical institutions that the MDC opens up," he says. "One of my long-term aims is to find out whether microRNAs might be competent for treating diseases."

Leaving New York is not easy, he says. "There is this special American spirit: everyone has the right to be happy. And Americans complain a lot less than Germans do."

Despite worries that Europe lags behind in genomics and systems biology, he is happy to be returning to Germany. "I want to bring home some of the things I learned in the United States," he says. "I feel like a 'European American'."

He and his wife, a pianist, are looking forward to the greater cultural variety of Europe. "Berlin in particular is a melting pot of science and culture," he says. He still plays the piano most mornings. "I feel healthy when I'm playing, in every sense."

Dirk Steuerwald

SCIENTISTS & SOCIETIES

Mass uprising of women in science

How do you revive a dormant chapter of a national organization? The Boston chapter of the Association of Women in Science (AWIS) had some unintentional help. A now-infamous speech last year by Harvard University president Larry Summers, calling women's scientific capabilities into question, brought in many new members.

AWIS was founded in 1971, dedicated to promoting women's full participation in all scientific fields through education, mentoring and job opportunities. AWIS also serves as a political voice for women in science, through testimonies on Capitol Hill and participation in a number of coalitions.

Membership of the local chapter MASS AWIS has risen above 80 individuals (including men), whose fields of interest span a broad range, including patent law, academic science, industrial science, engineering and education. In addition, our members range from undergraduates to retired scientists and are at various companies, hospitals, academic institutions and institutes. This diverse network, dedicated to ensuring a role for women in all scientific arenas, bridges the inter-institutional gaps and will provide role models for girls and women at all stages of their scientific careers. AWIS has the potential to play a key role in

promoting the advancement of women.

We are currently trying to organize a series of leadership workshops for undergraduates, graduates and postdocs that will target 'soft' skills, such as communication, leadership and interviewing. Although necessary to succeed in any career, these are not usually introduced in academic courses. MASS AWIS will also be establishing programmes and organizing mentors/volunteers to introduce the excitement of scientific discovery to school-age girls. Every chapter event serves as a valuable networking and mentoring resource for our members and guests. Finally, we are actively trying to build a scholarship fund to help disadvantaged women pursue their scientific dreams.

Besides spurring MASS AWIS membership, Summers provided a clear example of the continuing need for institutional change as well as ongoing education on gender equality. We all thought the effort to gain equality in scientific education and careers was well on its way, but this event and the statistics show that we are not yet at the top of the hill. Join an AWIS chapter and keep climbing with us.

Joanne Kamens is MASS AWIS chapter president, Karen Yee is MASS AWIS administrator.

▶ www.awis.org

GRADUATE JOURNAL

Selling the PhD

"Every day, do something that scares you," urges a song I heard once. Last month, I filled my quota for a long time. As part of my mentoring programme, I went for a mock job interview at a government agency. Questions ranged from the traditional strengths-and-weaknesses to testing my knowledge on politics.

In my answers, I tried to remember what people have told me about selling the PhD as a proof of your abilities rather than a scientific piece of work. Don't call yourself a researcher or scientist — rather say that you have experience in planning and realizing large projects. Don't say you're good at lab work — describe yourself as meticulous and patient. Beware of sounding arrogant. Sadly, one of the prevailing prejudices against PhDs is that they are stuck-up specialists. Show them you can think broadly and be a team player.

The interview was stressful but ultimately rewarding. Although my guess on a budget question was €35 million (US\$43 million) off, the interviewers said they liked my ability to think analytically and see the big picture, both products of my graduate education. It was a real boost to see I may be employable.

And the little budget mistake? Well, knowing how to find the facts is more important than actually knowing them by heart. At least, that's what I tell myself. If anybody else shares my opinion that €35 million don't matter, just send the money to me.

Katja Bargum is a graduate student in the Department of Biological and Environmental Sciences, University of Helsinki, Finland.

Directed energy

More than just a flash in the pan.

Jeff Hecht

A buzzing sound startled Ellen so she spilled her tea. After years of fighting the spread of insect-borne diseases in a warming world, bugs scared her more than her brother Bernie's war machines.

"Are you okay, Sis?" Bernie asked. "You look jittery."

"I thought I heard a mosquito." She had felt something tickling her left ear earlier, but it had only been the naturally shed macaw feather attached to her new earring.

He chuckled. "You can't be so jumpy in defence work."

She frowned older-sibling disapproval at her errant brother. "Controlling the spread of disease-carrying insects is a serious business, Bernie."

"So is defending against nuclear missiles."

"How many million people did they kill last year?"

"None, thanks to missile defence."

"Really, Bernie?" Ellen was eight years older, and after 52 years retained the wisdom of age. "How many nuclear missiles could your lasers really shoot down? You never even tested them."

"It's classified. I can't tell you."

That was his standard line. "It's classified because it doesn't work, and that's embarrassing, isn't it, Bernie?"

"It's a strategic defence system, Sis. It doesn't have to work every time. It's supposed to deter attack, so we only have to convince the other side that it could work."

"So it's all a bluff. I wish I could bluff mosquitoes away. Did anybody ever use anything you developed at Beltway Banditos?"

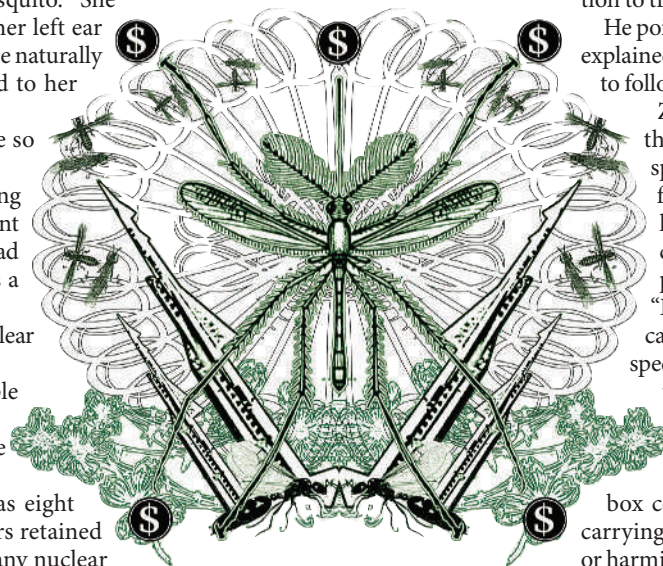
"The name was Beltway Systems," said Bernie, who had tired of the joke years ago. "Some of our technology has dual military and civilian uses. Remember that low-orbit satellite telephone system I told you about? The laser links between satellites used the pointing and tracking technology developed for our space-based laser interceptor."

"Did the phone satellites ever get off the ground?"

Bernie's face slumped, as it always did when one of his experiments went wrong, and he looked at the floor of his cluttered little house. "They laid me off before they finished the system, and I haven't checked."

Ellen signed. Bernie had spent 25 years in the defence industry, until the worsening climate crisis had finally drained the big money away from military hardware. "What are you working on now?"

Bernie's eyes brightened; he always liked to talk about his latest crazy idea. "I found more dual-use potential for the tracking and pointing technology. I've scaled it



down a bit, and I have a prototype out in the workshop. With your interest in bugs, you really should see it work."

"Where should I hide?" Ellen remembered the match-head rocket that had burned the family garage when Bernie was 14.

"You don't have to hide," Bernie said, rising from his chair. "You will need a pair of laser safety goggles, though."

When they reached his garage workshop, Bernie handed Ellen a pair of light pink wrap-around goggles. "They look pretty clear in the visible, but block the near-infrared beams almost completely."

The goggles made Bernie's hair and moustache look pink. "You're not going to burn holes in my clothes again, are you?" Ellen asked.

"No, nothing like that. It shouldn't harm clothing at all." Bernie put on a similar pair of goggles, pushed a wheeled red metal cabinet, then slid open the garage door.

Ellen saw the power cord to the cabinet was ominously thick. "Is this another of your crazy laser guns?"

"It can't sustain megawatt power, but it's good enough for insects."

"You've made a laser fly swatter?"

Bernie flipped open the hinged cover,

revealing a black metal bench cluttered with lenses and metal boxes. "This is to a fly swatter as the Internet is to a smoke signal. It includes sensors that track the position of each insect, processors to analyse their trajectories, a tracking and pointing system, and pulsed semiconductor lasers. It runs a modified version of software I developed for an orbiting laser battle station to track a fleet of nuclear warheads."

He pointed at a series of little boxes, and explained their functions too fast for Ellen to follow. Then he flipped a switch.

ZAP! A flash erupted in front of the box, and a little puff of smoke spread into the air. ZAP! Another flash. Except for his white hair and his middle-aged spread, Bernie could have been a 16-year-old playing with a new techno-toy. "Each zap takes out another bug. I can set it to pick out only certain species if you want, so it bags mosquitoes but not butterflies."

Ellen was speechless. Her brother had finally done something worthwhile. The little laser box could destroy incoming disease-carrying insects without using pesticides or harming beneficial types. She could see thousands of the boxes in village squares around the world, blocking the spread of disease. "Incredible!" she said. "Simply incredible."

Bernie glowed with pride.

"How did you manage to scale everything down?"

"Not quite everything," Bernie looked thoughtful for a moment. "It takes a lot of sensors and computer power to track the bugs, but it's a lot cheaper than missile defence."

"Wonderful," Ellen said, nodding her head. She saw one of the little laser boxes point toward her. A ZAP sizzled near her ear. "Ouch," she said, putting her hand up and feeling her earring. "Bernie! You burnt the feather off my earring!"

"It was moving; it must have fooled the sensors."

"That was a naturally shed feather from an endangered species. I paid \$70 for it!"

A familiar impish grin appeared on her brother's face. "We're still ahead. Each laser pulse costs only \$50."

Jeff Hecht writes regularly for *New Scientist* and *Laser Focus World*. His most recent book is *Beam: The Race to Make the Laser*. The match-head rocket he launched in his teens never set anything on fire.

JACEY